



University of Kentucky
UKnowledge

Theses and Dissertations--Mathematics

Mathematics

2013

RELATIVE PERTURBATION THEORY FOR DIAGONALLY DOMINANT MATRICES

Megan Dailey

University of Kentucky, meganmdailey@gmail.com

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Dailey, Megan, "RELATIVE PERTURBATION THEORY FOR DIAGONALLY DOMINANT MATRICES" (2013).
Theses and Dissertations--Mathematics. 11.
https://uknowledge.uky.edu/math_etds/11

This Doctoral Dissertation is brought to you for free and open access by the Mathematics at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Mathematics by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Megan Dailey, Student

Dr. Qiang Ye, Major Professor

Dr. Peter Perry, Director of Graduate Studies

RELATIVE PERTURBATION THEORY FOR DIAGONALLY DOMINANT
MATRICES

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Arts and Sciences
at the University of Kentucky

By
Megan Dailey
Lexington, Kentucky

Director: Dr. Qiang Ye, Professor of Mathematics
Lexington, Kentucky 2013

Copyright© Megan Dailey, 2013.

ABSTRACT OF DISSERTATION

RELATIVE PERTURBATION THEORY FOR DIAGONALLY DOMINANT MATRICES

Diagonally dominant matrices arise in many applications. In this work, we exploit the structure of diagonally dominant matrices to provide sharp entrywise relative perturbation bounds. We first generalize the results of Dopico and Koev to provide relative perturbation bounds for the LDU factorization with a well conditioned L factor. We then establish relative perturbation bounds for the inverse that are entrywise and independent of the condition number. This allows us to also present relative perturbation bounds for the linear system $Ax=b$ that are independent of the condition number. Lastly, we continue the work of Ye to provide relative perturbation bounds for the eigenvalues of symmetric indefinite matrices and non-symmetric matrices.

KEYWORDS: relative perturbation theory, relative error bounds, diagonally dominant matrices, LDU factorization, eigenvalues

Author's signature: _____ Megan Dailey

Date: _____ July 29, 2013

RELATIVE PERTURBATION THEORY FOR DIAGONALLY DOMINANT
MATRICES

By
Megan Dailey

Director of Dissertation: Qiang Ye

Director of Graduate Studies: Peter Perry

Date: July 29, 2013

This thesis is dedicated to Richie for showing me love and patience, even when I'm
at my crankiest.

ACKNOWLEDGMENTS

I would like to first thank my advisor Dr. Qiang Ye for the opportunity to work with him. I doubt I would be finishing now without your guidance and I truly appreciate all you have done for me. Secondly, I would like to thank Dr. Froilán Dopico for his collaboration and patience.

To the mathematics department at the University of Kentucky, I would like to thank you for your encouragement in and out of the classroom. Specifically to Dr. Peter Perry, who made it possible for me to teach at Centre College and gain valuable experience. I would also like to thank Sheri Rhine and Dr. David Leep, who offered invaluable help and advice during my job search. To the “Math Club” - Beth, Chris, Elizabeth, Dennis, and Nick - I would not have survived graduate school without your support and empathy. You all believed in me, even when I doubted myself.

I would also like to thank my friends from Centre College: Casey, Cindy, Crystal, Ginny, Jenn, Jenny, Joanna, Kate, Katie, Kristin, Rafayna, Rebecca, Stephanie, and Tallie. Your weekly emails reminded me there is a life outside of school and helped me focus on the bigger picture. I owe you all one big apology for all the venting.

Most of all, I would like to thank my family, especially my parents. Words can’t describe how much your love and support mean to me. Thank you for always being proud of me. I love you all.

TABLE OF CONTENTS

Acknowledgments	iii
Table of Contents	iv
Chapter 1 Introduction	1
1.1 Preliminaries and Notation	8
Chapter 2 LDU factorizations	12
2.1 Classical perturbation bounds	14
2.2 Relative perturbation bounds for diagonally dominant matrices	15
2.3 Generalized relative perturbation bounds for diagonally dominant matrices	21
Chapter 3 Inverses and Solutions to Linear Systems	45
3.1 Classical perturbation results	45
3.2 Relative perturbation results for diagonally dominant matrices	46
Chapter 4 Symmetric eigenvalue problem and singular value problem	54
4.1 Classical perturbation results	54
4.2 Relative perturbation results for diagonally dominant matrices	59
Chapter 5 Nonsymmetric Eigenvalue Problem	68
5.1 Classical perturbation results	70
5.2 Relative perturbation bounds	73
Chapter 6 Conclusions	80
Bibliography	81
Vita	85

Chapter 1 Introduction

Perturbation theory for a mathematical problem is the study of the effect small disturbances in the data have on the solution to the problem. The importance of perturbation theory is two-fold; perturbation theory allows us to investigate computational errors and approximate solutions to complicated systems as well as to study the stability of the solution.

Computations produced by numerical algorithms are plagued by two sources of error. One source of error is the data imputed into the algorithm. This error can be caused by prior calculations or from measurement errors. The second source of error is roundoff errors made within the algorithm itself. The standard method for analyzing roundoff errors is to use backward error analysis. In backward error analysis, we show that the computed solution produced by an algorithm is the exact solution to the problem with slightly perturbed inputs. In order to estimate the solution error, we must study how the solution of a problem is changed if we slightly alter, or perturb, the input data.

This thesis studies perturbation theory for linear algebra problems. The classical perturbation theory is extensively studied and surveyed in [41]. Consider the matrix eigenvalue problem, for example. Let A be a diagonalizable matrix and suppose $\tilde{A} = A + E$ is a perturbation of A . The classical perturbation result bounds the distance between an eigenvalue $\tilde{\lambda}$ of \tilde{A} and the closest eigenvalue λ of A as a multiple of the absolute perturbation E . That is,

$$|\tilde{\lambda} - \lambda| \leq \kappa(X)\|E\|_2 \tag{1.1}$$

where $A = X\Lambda X^{-1}$ with Λ a diagonal matrix and $\kappa(X) = \|X\|_2\|X^{-1}\|_2$. We denote by $\|X\|_2$ the spectral norm of X , which is the largest singular value of X .

The bound in (1.1) indicates absolute error but usually it is the relative error that is of interest. From (1.1), we have

$$\frac{|\tilde{\lambda} - \lambda|}{|\lambda|} \leq \kappa(X) \frac{\|E\|_2}{|\lambda|} = \kappa(X) \frac{\|A\|_2}{|\lambda|} \frac{\|E\|_2}{\|A\|_2} \quad (1.2)$$

The relative perturbation bound 1.2 that follows from (1.1) depends on the eigenvalue itself. This means that each eigenvalue has a different relative bound and those bounds for eigenvalues larger in magnitude are smaller than those that are smaller in magnitude. In the following example, we illustrate that this bound is as sharp as possible.

Example 1.1. *Let*

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

and consider the perturbed matrix

$$A + E = \begin{pmatrix} \lambda_1 + \epsilon & & \\ & \ddots & \\ & & \lambda_n + \epsilon \end{pmatrix}$$

where $\epsilon > 0$ is a small perturbation. The absolute perturbation bound (1.1) gives that error in an eigenvalue $\tilde{\lambda}$ of $A + E$ is bounded by

$$\min_i |\lambda_i - \tilde{\lambda}| \leq \|E\|_2 = \epsilon.$$

Suppose $\tilde{\lambda}$ is close to λ_{\min} , the eigenvalue of A of smallest magnitude. Then the relative error bound is

$$\frac{|\tilde{\lambda} - \lambda_{\min}|}{|\lambda_{\min}|} \leq \frac{\epsilon}{|\lambda_{\min}|}.$$

By definition of $A + E$, the relative error is $\epsilon/|\lambda_{\min}|$ and thus the bound presented in (1.2) is sharp.

In general, the bound (1.2) reflects the actual perturbation. However, in some cases, the eigenvalues are observed to exhibit much more relative accuracy than the classical theory would indicate. This becomes an issue when small eigenvalues are of interest because these error bounds indicate little relative accuracy. Consider the following example from [26].

Example 1.2. *Consider storing the matrix*

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

in floating point arithmetic. This produces the perturbed matrix

$$A + E = \begin{pmatrix} \lambda_1(1 + \epsilon_1) & & \\ & \ddots & \\ & & \lambda_n(1 + \epsilon_n) \end{pmatrix}$$

where $|\epsilon_i| \leq \epsilon$ and $\epsilon > 0$ reflects the machine accuracy. According to the absolute perturbation bound (1.1), the error in an eigenvalue $\tilde{\lambda}$ of $A + E$ is bounded by

$$\min_i |\lambda_i - \tilde{\lambda}| \leq \|E\|_2 = \max_k |\lambda_k \epsilon_k| \leq \epsilon \max_k |\lambda_k|.$$

For large eigenvalues, this bound is realistic. If $\tilde{\lambda}$ is near the eigenvalue of largest magnitude of A , λ_{max} , then the relative error is

$$\frac{|\lambda_{max} - \tilde{\lambda}|}{|\lambda_{max}|} \leq \epsilon.$$

However, if $\tilde{\lambda}$ is near the eigenvalue of smallest magnitude, λ_{min} , then the relative error is

$$\frac{|\lambda_{min} - \tilde{\lambda}|}{|\lambda_{min}|} \leq \epsilon \frac{|\lambda_{max}|}{|\lambda_{min}|}.$$

When the eigenvalues vary greatly in magnitude, this bound is much larger than ϵ . By the definition of $A + E$, the relative errors of all eigenvalues do not exceed ϵ and thus this bound is not tight.

One way to derive sharp relative perturbation bounds is to express the perturbation multiplicatively. That is, instead of representing the perturbation in A as $A + E$, we represent it as $D_1 A D_2$ where D_1 and D_2 are nonsingular matrices close to the identity matrix. The following is one such result.

Theorem 1.3 ([26]). *Let A be a diagonalizable matrix with eigenvalues $\{\lambda_i\}_{i=1}^n$ and let $\tilde{A} = D_1 A D_2$ where D_1 and D_2 are nonsingular. Let $\tilde{\lambda}$ be an eigenvalue of \tilde{A} , then*

$$\min_i |\lambda_i - \tilde{\lambda}| \leq |\tilde{\lambda}| \kappa(X) \|I - D_1^{-1} D_2^{-1}\|_2$$

In this theorem, by interpreting the perturbation as a multiplicative one, we can exploit certain matrix structures and perturbation structures to provide perturbation bounds for eigenvalues, even the smallest, with high relative accuracy. This is known as multiplicative perturbation theory, which has been extensively studied in [5, 10, 11, 12, 17, 18, 19, 29]. See [26] for an overview.

Although multiplicative perturbation theory provides nice relative perturbation bounds, a practical perturbation problem is generally concerned with an additive perturbation. On the other hand, if a matrix has certain structure, then one may express an additive perturbation as a multiplicative perturbation so as to derive relative perturbation bounds. One example where additive perturbations can easily be represented as multiplicative perturbations is the following.

Example 1.4. *Consider the symmetric tridiagonal matrix*

$$A = \begin{pmatrix} 0 & \alpha_1 & & & \\ \alpha_1 & 0 & \alpha_2 & & \\ & \alpha_2 & 0 & \alpha_3 & \\ & & \alpha_3 & 0 & \alpha_4 \\ & & & \alpha_4 & 0 & \alpha_5 \\ & & & & \alpha_5 & 0 \end{pmatrix}$$

and the component-wise relative perturbation

$$\tilde{A} = \begin{pmatrix} 0 & \tilde{\alpha}_1 & & & \\ \tilde{\alpha}_1 & 0 & \tilde{\alpha}_2 & & \\ & \tilde{\alpha}_2 & 0 & \tilde{\alpha}_3 & \\ & & \tilde{\alpha}_3 & 0 & \tilde{\alpha}_4 \\ & & & \tilde{\alpha}_4 & 0 & \tilde{\alpha}_5 \\ & & & & \tilde{\alpha}_5 & 0 \end{pmatrix}$$

where $\tilde{\alpha}_i = \alpha_i + e_i$, and $|e_i/\alpha_i| = |\epsilon_i| \leq \epsilon$. That is, $\tilde{\alpha}_i = \alpha_i(1 + \epsilon_i)$. Let $\beta_i = 1 + \epsilon_i$. Then, the perturbed matrix \tilde{A} can be represented as a multiplicative perturbation $\tilde{A} = DAD$, where

$$D = \begin{pmatrix} \beta_1 & & & & \\ & 1 & & & \\ & & \beta_2 & & \\ & & & \frac{\beta_3}{\beta_2} & \\ & & & & \frac{\beta_4\beta_2}{\beta_3} \\ & & & & & \frac{\beta_5\beta_3}{\beta_4\beta_2} \end{pmatrix}$$

Using the multiplicative relative perturbation in Theorem 1.3 we obtain

$$\min_i \frac{|\lambda_i - \tilde{\lambda}|}{|\tilde{\lambda}|} \leq 1 - \left(\frac{1 - \epsilon}{1 + \epsilon} \right)^4 = 8\epsilon + \mathcal{O}(\epsilon^2).$$

On the other hand, if treating \tilde{A} as an additive perturbation, we have

$$\min_i \frac{|\lambda_i - \tilde{\lambda}|}{|\lambda_i|} \leq \frac{2\epsilon}{|\lambda_i|}.$$

Perturbations that can be easily written multiplicatively do not commonly arise in practice. In general, we are interested in additive perturbations. Thus, it is important to study perturbation theory in terms of the perturbed elements by exploiting specific structures.

In this thesis, we will consider an important class of matrices, i.e. diagonally dominant matrices, for which strong relative perturbation bounds can be derived.

A matrix is called (row) diagonally dominant if for each row, the diagonal entry is larger than the sum of the absolute value of the off-diagonal entries. Diagonally dominant matrices arise in many applications and have been extensively studied, see [2, 1, 3, 13, 14, 15, 40, 43, 44]. For instance, they appear frequently in numerical solutions for both ordinary and partial differential equations. Diagonally dominant matrices exhibit very nice theoretical properties, as explained in [24, 25]. They are a class of matrices for which iterative solution methods can be successfully applied.

The structure of diagonally dominant matrices has been exploited to provide strong relative perturbation bounds. In [44], a perturbation bound is presented for the eigenvalues of a symmetric positive semi-definite diagonally dominant matrix that improves upon previous research in that the bound is independent of any condition number. In [43], an algorithm is presented for computing the singular values of diagonally dominant matrices with relative errors on the order of machine precision. For symmetric positive semi-definite diagonally dominant matrices, i.e. diagonally dominant matrices with positive diagonal entries, this algorithm also computes the eigenvalues to the order of machine precision. In [13], a structured perturbation theory for the LDU factorization of diagonally dominant matrices is presented. $A = LDU$ is an LDU factorization if L is a unit lower triangular matrix, D is a diagonal matrix, and U is a unit upper triangular matrix. The relative errors for the diagonal matrix D are bounded componentwise and the relative errors for the factors L and U are norm-wise. In [16], sharp relative perturbation bounds for solutions to linear systems are derived using a structured perturbation in a rank revealing factorization such as the LDU factorization.

In this work we systematically derive relative perturbation bounds for a variety of linear algebra problems for diagonally dominant matrices. We first generalize the perturbation bounds presented by Dopico and Koev [13] for the LDU factorization of diagonally dominant matrices. The bounds presented in [13] require complete di-

agonal pivoting which does not ensure a well-conditioned factor L . We prove that by using the so called column diagonal dominance pivoting, we can produce relative perturbation bounds for the LDU factorization that ensures L is column diagonally dominant and hence well-conditioned. Specifically, for a row diagonally dominant matrix A that is not entirely zero, there is at least one k such that $a_{kk} \neq 0$ and column k is column diagonally dominant. The column diagonal dominance pivoting strategy is to permute row 1 with row k and column 1 with column k , after which the first column of the matrix is diagonally dominant. If there are many columns that are column diagonally dominant, we choose the one with maximal a_{kk} . Applying Gaussian elimination, the first column of L that is produced is column diagonally dominant. Repeating this strategy, at the end of Gaussian elimination we obtain a row diagonally dominant U as usual, but now L is column diagonally dominant. Using [39], we have that the condition numbers of L and U are bounded as $\kappa_\infty(L) \leq n^2$ and $\kappa_\infty(U) \leq 2n$. Having a well-conditioned L is critical in many applications of the LDU factorization. We then, for diagonally dominant matrices, obtain relative perturbation bounds for several other linear algebra problems including inverses, solutions to linear systems, eigenvalues, and singular values. These new results are being prepared for publication [7, 8].

This thesis is organized as follows: In Chapter 2, we focus on the LDU factorization. We discuss the classical results in literature and present the current relative perturbation results of Dopico and Koev [13]. We then present the main result of this thesis that provides a stronger form of the LDU perturbation bounds from [13]. The proof of the new bound involves some significantly new techniques. In Chapter 3 we discuss perturbation theory for the matrix inverse and solution to linear systems. After presenting the classical results, we present relative perturbation results for diagonally dominant matrices. The symmetric eigenvalue problem is the focus of Chapter 4. We present classical additive and relative perturbation bounds for sym-

metric matrices, see [26]. We also include in this chapter results for the singular value problem. In Chapter 5, we consider the nonsymmetric eigenvalue problem.

1.1 Preliminaries and Notation

Rank-revealing decompositions have been a key component to finding high accuracy singular value decompositions (SVDs). In [10], Demmel et. al produce an algorithm to compute high accuracy SVDs by first computing any rank revealing decomposition, i.e. $A = XDY^T$, then the SVD of XDY^T is computed using an algorithm of Jacobi type. In general, we say a decomposition $A = XDY^T$ is a *rank revealing decomposition* if X and Y are well conditioned and D is diagonal and nonsingular. The criteria “well conditioned” is dependent on the problem at hand and the desired error tolerance. The SVD itself and the decomposition that results from Gaussian Elimination with complete pivoting are examples of rank-revealing decompositions.

Let $A = XDY^T$ and $\tilde{A} = \tilde{X}\tilde{D}\tilde{Y}^T$ with $D = \text{diag}(d_i)$ and $\tilde{D} = \text{diag}(\tilde{d}_i)$. In [10], it is shown that if \tilde{A} is a perturbation of A satisfying

$$|\tilde{d}_i - d_i| \leq \epsilon |d_i|$$

$$\|\tilde{X} - X\|_2 \leq \epsilon \|X\|_2, \quad \|\tilde{Y} - Y\|_2 \leq \epsilon \|Y\|_2$$

where $0 \leq \epsilon < 1$, then, the singular values of A and \tilde{A} satisfy

$$\frac{|\tilde{\sigma}_i - \sigma_i|}{\sigma_i} \leq 2\eta + \eta^2$$

where $\eta = \epsilon(2 + \epsilon) \max\{\kappa(X), \kappa(Y)\}$, with $\kappa(Z) = \|Z\|_2 \|Z^{-1}\|_2$ the condition number of Z defined in terms of the Moore-Penrose pseudoinverse.

Note that the error in the singular values is dependent on the condition number of the factors X and Y , which are assumed to be well-conditioned, and not the condition number of the matrix A , which could be large. Rank revealing decompositions are also utilized in [14] and [15] to bound perturbation errors in eigenvalues of symmetric

matrices with high relative accuracy. In [16], accurate rank revealing decompositions are used to compute accurate solutions to structured linear systems.

Recent research has focused on special matrices for which eigenvalues and singular values can be computed with high relative accuracy. Among these matrices are diagonally dominant matrices, which we discuss now.

Definition 1.5. *A matrix $A = [a_{ij}]$ is said to be (row) diagonally dominant if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$ for all i and is said to be column diagonally dominant if $|a_{ii}| \geq \sum_{j \neq i} |a_{ji}|$ for all i .*

An idea that has played a key role in computing eigenvalues of diagonally dominant matrices is to reparameterize the matrix in terms of its diagonally dominant parts and off diagonal entries. This reparameterization is introduced in [2, 43] and stated below.

Definition 1.6. *Given an $n \times n$ matrix $M = [m_{ij}]$ and an n -vector $v = [v_i]$, we use $\mathcal{D}(M, v)$ to denote the matrix $A = [a_{ij}]$ whose off-diagonal entries are the same as M and whose i th diagonal entry is $a_{ii} = v_i + \sum_{j \neq i} |m_{ij}|$. Namely, we write*

$$A = \mathcal{D}(M, v)$$

and call it the representation of A by diagonally dominant parts v if

$$a_{ij} = m_{ij} \text{ for } i \neq j; \text{ and } a_{ii} = v_i + \sum_{j \neq i} |m_{ij}|$$

Given a matrix $A = [a_{ij}]$, we denote by A_D the matrix whose off-diagonal entries are the same as A and whose diagonal entries are zero. Then, letting $v_i = a_{ii} - \sum_{j \neq i} |a_{ij}|$ and $v = (v_1, v_2, \dots, v_n)^T$, we have

$$A = \mathcal{D}(A_D, v)$$

as the representation of A by diagonally dominant parts.

Theorem 1.7 ([13]). *If $A \in \mathbf{R}^{n \times n}$ is diagonally dominant, then*

1. *Every principal submatrix of A is diagonally dominant;*
2. *PAP^T is diagonally dominant for every permutation matrix $P \in \mathbf{R}^{n \times n}$;*
3. *If $a_{11} \neq 0$ then the Schur complement of a_{11} in A is diagonally dominant;*
4. *If $\det A \neq 0$ then $\det A$ has the same sign as the product $a_{11}a_{22} \cdots a_{nn}$; and*
5. *$|\det A(i', i')| \geq |\det A(i', j')|$, for all $i = 1, \dots, n$ and for all $j \neq i$ where $A(i', j')$ denotes the submatrix of A formed by deleting row i and column j of A .*

Diagonally dominant matrices have several nice properties that can be utilized, see [13, Theorem 1]. For instance, strictly diagonally dominant matrices (that is, diagonally dominant matrices as defined in Definition 1.5 with strict inequalities) are nonsingular and Gaussian elimination can be performed without interchanging rows. This implies that a strictly diagonally dominant matrix A has an LDU factorization. That is, that we can write A as $A = LDU$ where L , D , and U are a lower triangular matrix, diagonal matrix, and upper triangular matrix, respectively.

Notation: In this thesis we consider only real matrices and we denote by $\mathbf{R}^{m \times n}$ the set of $m \times n$ real matrices. The entries of matrix A are a_{ij} . We use MATLAB notation for submatrices. We use $i : j$ to denote the index subset from i to j . That is, $A(i : j, k : l)$ denotes the submatrix of A formed by rows i through j and columns k through l . We use $A(i', j')$ to denote the submatrix of A formed by deleting row i and column j from A . Let $\alpha = [i_1, i_2, \dots, i_q]$ where $1 \leq i_1 < i_2 < \dots < i_q \leq n$. Then $A(\alpha, \alpha)$ denotes the submatrix of A that consists of rows i_1, i_2, \dots, i_q and columns i_1, i_2, \dots, i_q . We denote by $\|\cdot\|$ a general matrix operator norm. We will use five special matrix norms: the max norm $\|A\|_{\max} = \max_{ij} |a_{ij}|$, the maximum absolute column sum $\|A\|_1 = \max_j \sum_i |a_{ij}|$, the maximum absolute row sum $\|A\|_\infty = \max_i \sum_j |a_{ij}|$, the spectral norm $\|A\|_2$, which is the largest singular value of A , and the Frobenius

norm $\|A\|_F = (\sum_{i,j} |a_{ij}|^2)^{1/2}$. The transpose of A is denoted by A^T . We denote by A^* the conjugate transpose of A . The sign of $x \in \mathbf{R}$ is $\text{sign}(x)$, where $\text{sign}(0)$ is defined to be 1.

Chapter 2 LDU factorizations

In this chapter we focus on one of the most important matrix factorizations in numerical analysis. The LU factorizations decomposes a matrix into the product of a unit lower triangular matrix L and an upper triangular matrix U . The LU factorization is primarily used to solve systems of linear equations, inverting a matrix, and computing the determinant.

In regards to solving systems of linear systems, the focus is usually on backward error analysis. Consider computing the LU factorization of A without pivoting. The computed factors \tilde{L} and \tilde{U} are the exact L and U factors of some matrix $A + E$. Backward error analysis provides a bound for the perturbation E .

Theorem 2.1 ([9]). *Let A be a matrix and suppose \tilde{L} and \tilde{U} are the computed factors in the LU decomposition. Then, $\tilde{L}\tilde{U} = A + E$ with*

$$|E| \leq (n\epsilon + \mathcal{O}(\epsilon^2))|\tilde{L}||\tilde{U}|$$

where $\epsilon > 0$ is machine precision.

The LU factorization is used to solve the linear system $Ax = b$, or rather $LUx = b$, by first solving $Ly = b$ for y and then $Ux = y$ for x . Each of these can be easily solved by substitution. From Theorem 2.1 we obtain the following result for the backward error analysis on the solution to $Ax = b$.

Corollary 2.2 ([9]). *Let \tilde{x} be the computed solution of $Ax = b$ by using the computed factors \tilde{L} and \tilde{U} . Then, \tilde{x} is the exact solution to $(A + E)\tilde{x} = b$ with*

$$\|E\|_{\infty} \leq 3g_{pp}n^3\epsilon\|A\|_{\infty}$$

where $g_{pp} = \frac{\|\tilde{U}\|_{\max}}{\|A\|_{\max}}$.

From backward error analysis, we can derive forward error bounds using perturbation analysis. For example, to derive forward error bounds for the LU factorization, we develop a perturbation theory for the L and U factors of a perturbed matrix A . Barrlund [4] presented a normwise bound and Sun [21] presented a componentwise bound for the factors of the LU decomposition of a general matrix. Dopico and Bueno [6] improve these results for the special case of tridiagonal matrices without pivoting. In Section 2.1, we will examine the strong relative perturbation bounds of Dopico and Koev [13] for the LDU factorization of diagonally dominant matrices with nonnegative diagonals under a structured perturbation. We will improve upon these results by generalizing the structured perturbation in Section 2.2.

Definition 2.3. *A matrix $A \in \mathbf{R}^{n \times n}$ with rank r is said to have LDU factorization if there exists a unit lower triangular matrix $L_{11} \in \mathbf{R}^{r \times r}$, a unit upper triangular matrix $U_{11} \in \mathbf{R}^{r \times r}$, and a nonsingular diagonal matrix $D_{11} \in \mathbf{R}^{r \times r}$ such that $A = LDU$ where*

$$L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & I_{n-r} \end{bmatrix}, \quad D = \begin{bmatrix} D_{11} & 0 \\ 0 & 0_{n-r} \end{bmatrix}, \quad U = \begin{bmatrix} U_{11} & U_{12} \\ 0 & I_{n-r} \end{bmatrix}$$

Note that the LDU factorization is another form of the LU factorization. Namely, from $A = LDU$ above, $A = LU_1$ with $U_1 = DU$ is the standard LU factorization. On the other hand, from $A = LU$, $A = LDU_1$ with $U_1 = D^{-1}U$ is the LDU factorization as defined above assuming A is invertible. For a general diagonally dominant matrix, applying any pivoting scheme with simultaneous row and column permutations leads to a matrix that has LDU factorization.

Theorem 2.4 ([20]). *If A has LDU factorization, then the factorization is unique*

and the nontrivial entries of L , D , and U are given by

$$l_{ij} = \frac{\det A([1 : j-1, i], 1 : j)}{\det A(1 : j, 1 : j)}, \quad i > j \text{ and } j = 1, \dots, r, \quad (2.1)$$

$$d_i = \frac{\det A(1 : i, 1 : i)}{\det A(1 : i-1, 1 : i-1)}, \quad i = 1, \dots, r, (\det A(1 : 0, 1 : 0) := 1) \quad (2.2)$$

$$u_{ij} = \frac{\det A(1 : i, [1 : i-1, j])}{\det A(1 : i, 1 : i)}, \quad i < j \text{ and } i = 1, \dots, r. \quad (2.3)$$

where $[1 : j-1, i]$ denotes the index subset containing indices 1 to $j-1$ and i .

2.1 Classical perturbation bounds

The normwise and element-wise bounds of Barrlund [4] and Sun [21] are combined in [23] and presented below.

Theorem 2.5. *Let the nonsingular matrices $A \in \mathbf{R}^{n \times n}$ and $A + \Delta A$ have LU factorizations $A = LU$ and $A + \Delta A = (L + \Delta L)(U + \Delta U)$, and assume that $\|G\|_2 < 1$ where $G = L^{-1}\Delta AU^{-1}$. Then*

$$\max \left\{ \frac{\|\Delta L\|_F}{\|L\|_2}, \frac{\|\Delta U\|_F}{\|U\|_2} \right\} \leq \frac{\|G\|_F}{1 - \|G\|_2} \leq \frac{\|L^{-1}\|_2 \|U^{-1}\|_2 \|A\|_2}{1 - \|L^{-1}\|_2 \|U^{-1}\|_2 \|\Delta A\|_2} \frac{\|\Delta A\|_F}{\|A\|_F} \quad (2.4)$$

Moreover, if $\rho(|\tilde{G}|) < 1$, where $\tilde{G} = (L + \Delta L)^{-1}\Delta A(U + \Delta U)^{-1}$, then

$$\begin{aligned} |\Delta L| &\leq |L + \Delta L| \text{stril} \left((I - |\tilde{G}|)^{-1} |\tilde{G}| \right), \\ |\Delta U| &\leq \text{triu} \left(|\tilde{G}| (I - |\tilde{G}|)^{-1} \right) |U + \Delta U|, \end{aligned}$$

where $\text{stril}(\cdot)$ and $\text{triu}(\cdot)$ denote, respectively, the strictly lower triangular part and the upper triangular part of their matrix arguments.

The term $\chi(A) := \|L^{-1}\|_2 \|U^{-1}\|_2 \|A\|_2$ in the normwise bounds in Theorem 2.5 serves as an upper bound for the condition number for the LU factorization of A . Note that $\chi(A)$ is larger than the traditional condition number $\kappa(A)$. A stronger bound would use $\kappa(A)$ instead of $\chi(A)$. Another disadvantage to the bounds on ΔL and ΔU in Theorem 2.5 is that they include the factors L and U themselves. Simpler

bounds can be achieved by focusing on specific structures. In the next chapter we are able to significantly improve this result for diagonally dominant matrices in two ways. First, it is independent of any condition number. Second, it is an entrywise relative bound for D (or the diagonal of U) and is independent of A and the factors L and U . In fact, it is only dependent upon the size of the matrix and the perturbation.

2.2 Relative perturbation bounds for diagonally dominant matrices

In this section, we will focus on the perturbation results for the LDU factorization of diagonally dominant matrices with nonnegative diagonals provided by Dopico and Koev in [13]. Their result is based on a diagonal pivoting scheme.

In general, consider applying the Gaussian elimination to A with a pivoting strategy. Assuming that k steps of elimination have been performed, we let $A^{(k+1)} = [a_{ij}^{(k+1)}]$ denote the matrix after the k -th Gaussian elimination, and we write $A^{(1)} = A$. It is well known that $A^{(k)}$ is diagonally dominant if A is diagonally dominant. We represent it as $A^{(k)} = \mathcal{D}(A_D^{(k)}, v_D^{(k)})$ where $v^{(k)} = [v_1^{(k)}, v_2^{(k)}, \dots, v_n^{(k)}]^T$ with

$$v_i^{(k)} = a_{ii}^{(k)} - \sum_{j=k, j \neq i}^n |a_{ij}^{(k)}| \quad (2.5)$$

for $i \geq k$ and $v_i^{(k)} = v_i^{(k-1)}$ for $i < k$. Let r be the rank of A . Then, we can perform r steps of Gaussian elimination to produce the matrix $A^{(r+1)}$. Determinantal formulas for the entries of $A^{(k)}$ will be of interest. From [20], we have

$$a_{ij}^{(k+1)} = \frac{\det A([1 : k, i], [1 : k, j])}{\det A(1 : k, 1 : k)} \quad (2.6)$$

for $k + 1 \leq i, j \leq n$ and $1 \leq k \leq \min\{r, n - 1\}$.

Definition 2.6. A diagonally dominant matrix $A \in \mathbf{R}^{n \times n}$ with rank r is said to be arranged for complete-diagonal pivoting if

$$|a_{kk}^{(k)}| = \max_{k \leq i \leq n} |a_{ii}^{(k)}|, \quad k = 1, \dots, \min\{r, n - 1\}.$$

In general, there exists a permutation P so that PAP^T is arranged for complete-diagonal pivoting.

Theorem 2.7. *Let $A = \mathcal{D}(A_D, v) \in \mathbb{R}^{n \times n}$ be such that $v \geq 0$. Suppose A has LDU factorization $A = LDU$ as in Definition 2.3. Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbb{R}^{n \times n}$ be such that*

$$|\tilde{v} - v| \leq \epsilon v \text{ and } |\tilde{A}_D - A_D| \leq \epsilon |A_D|, \text{ for some } 0 \leq \epsilon < 1. \quad (2.7)$$

Then,

1. \tilde{A} is row diagonally dominant with nonnegative diagonal entries and has LDU factorization $\tilde{A} = \tilde{L}\tilde{D}\tilde{U}$ where $\tilde{L} = [\tilde{l}_{ij}]$, $\tilde{D} = [\tilde{d}_{ij}]$, and $\tilde{U} = [\tilde{u}_{ij}]$;

2. For $i = 1, \dots, n$,

$$|\tilde{d}_{ii} - d_{ii}| \leq \frac{2n\epsilon}{1 - 2n\epsilon} |d_{ii}|,$$

3. for $i < j$,

$$|\tilde{u}_{ij} - u_{ij}| \leq 3n\epsilon,$$

and

$$\frac{\|\tilde{U} - U\|_\infty}{\|U\|_\infty} \leq 3n^2\epsilon$$

4. and, if A is arranged for complete diagonal pivoting, for $i > j$

$$|\tilde{l}_{ij} - l_{ij}| \leq \frac{3n\epsilon}{1 - 2n\epsilon}$$

and

$$\frac{\|\tilde{L} - L\|_\infty}{\|L\|_\infty} \leq \frac{3n^2\epsilon}{1 - 2n\epsilon}$$

Theorem 2.7 states that small perturbations in the diagonally dominant parts and the off diagonal entries, lead to small relative perturbations in the entries of D , small absolute perturbations in the entries of U , and if we assume the matrix is arranged for complete diagonal pivoting, small absolute perturbations in the entries of L which also imply small normwise relative perturbation bounds in L and U . At each stage of

Gaussian elimination with complete diagonal pivoting, the same row and column are exchanged to place in the pivot position the diagonal entry with the largest absolute value of the corresponding Schur complement.

In the LDU factorization with complete diagonal pivoting, the factor U is row diagonally dominant, and hence well conditioned. However, the factor L does not inherit, in general, any particular property. The following example from [13] illustrates the necessity of a pivoting scheme to guarantee good behavior in the L factor under structured perturbation of type (2.7).

Example 2.8. *Consider the LDU factorization of the following diagonally dominant matrix $A = LDU$ without pivoting*

$$A = \begin{bmatrix} 1000 & 100 & 500 \\ 0 & 0.1 & 0.05 \\ 100 & 10 & 120 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0.1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1000 & & \\ & 0.1 & \\ & & 70 \end{bmatrix} \begin{bmatrix} 1 & 0.1 & 0.5 \\ & 1 & 0.5 \\ & & 1 \end{bmatrix}.$$

Observe that the vector of diagonally dominant parts of A is $v = [400, 0.05, 10]$. Now consider the LDU factorization of the diagonally dominant matrix $\tilde{A} = \tilde{L}\tilde{D}\tilde{U}$ without pivoting

$$\begin{aligned} \tilde{A} &= \begin{bmatrix} 1000 & 101 & 500 \\ 0 & 0.1 & 0.05 \\ 100 & 10 & 120 \end{bmatrix} \\ &= \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0.1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1000 & & \\ & 0.1 & \\ & & 70.05 \end{bmatrix} \begin{bmatrix} 1 & 0.101 & 0.5 \\ & 1 & 0.5 \\ & & 1 \end{bmatrix} \end{aligned}$$

whose vector of diagonally dominant parts is $\tilde{v} = [399, 0.05, 10]$. Note that A and \tilde{A} satisfy (2.7) with $\epsilon = 10^{-2}$ but their L factors are very different since $|\tilde{l}_{32} - l_{32}| = 1$.

Suppose P is a permutation matrix such that PAP^T is arranged for complete

diagonal pivoting. Let $B = PAP^T$ and consider its LDU factorization

$$B = \begin{bmatrix} 1000 & 500 & 100 \\ 100 & 120 & 10 \\ 0 & 0.05 & 0.1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0.1 & 1 & \\ 0 & \frac{0.05}{70} & 1 \end{bmatrix} \begin{bmatrix} 1000 & & \\ & 70 & \\ & & .1 \end{bmatrix} \begin{bmatrix} 1 & 0.5 & 0.1 \\ & 1 & 0 \\ & & 1 \end{bmatrix}.$$

Let $\tilde{B} = P\tilde{A}P^T$ and consider its LDU factorization

$$\begin{aligned} \tilde{B} &= \begin{bmatrix} 1000 & 500 & 100 \\ 100 & 120 & 10 \\ 0 & 0.05 & 0.1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & & \\ 0.1 & 1 & \\ 0 & \frac{0.05}{140} & 1 \end{bmatrix} \begin{bmatrix} 1000 & & \\ & 70 & \\ & & .1(1 + \frac{0.05}{70}) \end{bmatrix} \begin{bmatrix} 1 & 0.5 & 0.101 \\ & 1 & \frac{-0.1}{70} \\ & & 1 \end{bmatrix}. \end{aligned}$$

Notice that the LDU factorizations for B and \tilde{B} are very close to each other.

In the next section, we introduce a different pivoting strategy that will ensure L is column diagonally dominant and hence well conditioned. Moreover, it still produces small relative perturbation in the LDU factorization. For the rest of this section, we present several results from [13] which will be used in proving new bounds.

The proof of Theorem 2.7 hinges on Theorem 2.4 and several perturbation results for determinants. It is useful to consider the cofactor expansion of the determinant of a diagonally dominant matrix, as given in Lemma 2.9 below.

Lemma 2.9 ([13]). *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$. Denote the algebraic cofactors of A by*

$$C_{ij} := (-1)^{i+j} \det A(i', j'), \quad i, j = 1, \dots, n.$$

Then

$$\det A = v_i C_{ii} + \sum_{j \neq i} (|a_{ij}| C_{ii} + a_{ij} C_{ij}), \quad i = 1, \dots, n.$$

with $v_i C_{ii} \geq 0$ and $(|a_{ij}| C_{ii} + a_{ij} C_{ij}) \geq 0$ for $j \neq i$.

In Lemma 2.10, Dopico and Koev present perturbation results for the determinant of diagonally dominant matrices with nonnegative diagonals under structured perturbations of type (2.7).

Lemma 2.10 ([13]). *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ satisfy (2.7). Suppose $\tilde{A}^{[i]} = \mathcal{D}(\tilde{A}_D^{[i]}, \tilde{v}^{[i]}) \in \mathbf{R}^{n \times n}$ is a matrix that differs from A in only the i th row and whose i th row is the same as the i th row of \tilde{A} . Then*

$$\det \tilde{A}^{[i]} = (\det A)(1 + \eta_i), \text{ where } |\eta_i| \leq \epsilon. \quad (2.8)$$

This result is then utilized to provide a perturbation result for principal minors of diagonally dominant matrices with nonnegative diagonals.

Lemma 2.11 ([13]). *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ satisfy (2.7). Suppose $\tilde{A}^{[i]} = \mathcal{D}(\tilde{A}_D^{[i]}, \tilde{v}^{[i]}) \in \mathbf{R}^{n \times n}$ is a matrix that differs from A in only the i th row and whose i th row is the same as the i th row of \tilde{A} . Let $1 \leq i_1 < i_2 < \dots < i_q \leq n$ and $\alpha = [i_1, i_2, \dots, i_q]$, and denote the principle submatrix of A that lies in rows and columns indexed by α as $A(\alpha, \alpha)$. Then we have*

$$\det \tilde{A}^{[i]}(\alpha, \alpha) = \begin{cases} \det A(\alpha, \alpha) & \text{if } i \notin \alpha \\ (\det A(\alpha, \alpha))(1 + \delta_i^{(\alpha)}) & \text{if } i \in \alpha \end{cases} \quad (2.9)$$

where $|\delta_i^{(\alpha)}| \leq \epsilon$ and

$$\det \tilde{A}(\alpha, \alpha) = (\det A(\alpha, \alpha))(1 + \eta_1^{(\alpha)}) \cdots (1 + \eta_q^{(\alpha)}) \quad (2.10)$$

where $|\eta_k^{(\alpha)}| \leq \epsilon$ for $k = 1, \dots, q$.

Consider the determinants in Theorem 2.4. In [13], the following notation is introduced for simplicity

$$g_{pq}^{(k+1)} = \det A([1 : k, p], [1 : k, q]), \quad (2.11)$$

for $1 \leq k \leq n-1$ and $k-1 \leq p, q \leq n$. We denote by $(\tilde{g}^{[i]})_{pq}^{(k+1)}$ and $\tilde{g}_{pq}^{(k+1)}$ the corresponding minors of the perturbed matrices $\tilde{A}^{[i]}$ and \tilde{A} as defined in Lemmas 2.10 and 2.11 above, respectively. Lemma 2.12 establishes relationships that are utilized in Lemma 2.13 to provide perturbation results for nonprinciple minors.

Lemma 2.12 ([13]). *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$. For $k = 1, \dots, n-2$, $p \neq q$, and $k+1 \leq p, q \leq n$, let G_{ij} be the algebraic cofactor of $A([1:k, p], [1:k, q])$ for the entry a_{ij} . Then, for the minors defined in (2.11),*

$$g_{pq}^{(k+1)} = a_{p1}G_{p1} + \dots + a_{pk}G_{pk} + a_{pq}G_{pq} \quad (2.12)$$

$$2g_{pp}^{(k+1)} \geq |a_{p1}G_{p1}| + \dots + |a_{pk}G_{pk}| + |a_{pq}G_{pq}| \quad (2.13)$$

and for $1 \leq i \leq k$,

$$g_{pq}^{(k+1)} = \left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right) G_{ii} + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} (a_{ij}G_{ij} + |a_{ij}|G_{ii}) \quad (2.14)$$

$$2g_{pp}^{(k+1)} \geq \left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right) |G_{ii}| + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij}G_{ij} + |a_{ij}|G_{ii}| \quad (2.15)$$

Lemma 2.13 ([13]). *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $1 \leq k \leq n-2$, $k+1 \leq p, q \leq n$ and $p \neq q$. Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ that satisfies (2.7) with the given p . Suppose $\tilde{A}^{[i]} = \mathcal{D}(\tilde{A}_D^{[i]}, \tilde{v}^{[i]}) \in \mathbf{R}^{n \times n}$ is a matrix that differs from A in only the i th row and whose i th row parameters are the entries in the i th row of \tilde{A} . Then, the following statements hold*

1. $\left| (\tilde{g}^{[i]})_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \leq \begin{cases} 0 & \text{if } i \notin [1:k, p] \\ 2g_{pp}^{(k+1)} & \text{if } i \in [1:k, p] \end{cases}$
2. $|\tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)}| \leq 2((1+\epsilon)^{k+1} - 1)g_{pp}^{(k+1)}.$

2.3 Generalized relative perturbation bounds for diagonally dominant matrices

The LDU perturbation results of Dopico and Koev [13] require complete diagonal pivoting which does not guarantee a well conditioned factor L . In this section, we utilize a different pivoting scheme that ensures L is column diagonally dominant. For many applications of the LDU factorization, it is desirable to have both L and U well-conditioned. For example, a LDU factorization with well-conditioned L and U is a rank revealing factorization which can be used to accurately compute singular values [43], among many other uses. In later chapters, we use the LDU factorization to derive new relative perturbation bounds for other problems, some of which also rely on well-conditioned L factor.

In [43], a pivoting strategy, previously suggested by Pena [39], is presented that produces a column diagonally dominant L . This strategy is referred to as column diagonal dominance pivoting. At each stage of Gaussian elimination, the same row and column are exchanged to place in the pivot position the maximal diagonal entry that is column diagonally dominant. That is, at step k , we exchange the same row and column such that

$$a_{kk}^{(k)} = \max_{k \leq i \leq n} \{a_{ii}^{(k)} : a_{ii}^{(k)} - \sum_{j=k, j \neq i}^n |a_{ji}^{(k)}| \geq 0\}$$

Note that for a row diagonally dominant matrix, there is at least one column that is diagonally dominant and thus the set $\{a_{ii}^{(k)} : a_{ii}^{(k)} - \sum_{j=k, j \neq i}^n |a_{ji}^{(k)}| \geq 0\}$ is not empty. At the end, we obtain a row diagonally dominant U as usual, but now L is column diagonally dominant. Hence, by [39]

$$\|L\|_{\infty} \leq n \quad \|L^{-1}\|_{\infty} \leq n \quad (2.16)$$

$$\|U\|_{\infty} \leq 2 \quad \|U^{-1}\|_{\infty} \leq n \quad (2.17)$$

which implies the condition numbers of L and U are bounded as $\kappa_{\infty}(L) \leq n^2$ and $\kappa_{\infty}(U) \leq 2n$. From here on we will assume the matrix A is already permuted under

the column diagonal dominance pivoting scheme. In particular, the use of column diagonal dominance pivoting allows us to bound the column sums of the L factor in terms of the diagonally dominant part and the diagonal entry of A as seen in the following theorem.

Theorem 2.14. *Let $r = \text{rank}(A)$. For $k \leq r$, we have*

$$\sum_{i=k+1}^n |a_{ik}^{(k)}| + v_i^{(k)} \leq (n-k)a_{kk}^{(k)}$$

Proof. At step k , define $\delta_i^{(k)} = a_{ii}^{(k)} - \sum_{j=k, j \neq i}^n |a_{ji}^{(k)}|$. Then, we have

$$a_{kk}^{(k)} = \max\{a_{ii}^{(k)} : \delta_i^{(k)} \geq 0, i \geq k\}.$$

If all $\delta_i^{(k)} \geq 0$, then

$$\begin{aligned} \sum_{i=k+1}^n (|a_{ik}^{(k)}| + v_i^{(k)}) &\leq \sum_{i=k+1}^n a_{ii}^{(k)} \\ &\leq \sum_{i=k+1}^n a_{kk}^{(k)} \leq (n-k)a_{kk}^{(k)} \end{aligned}$$

Otherwise, if there is at least one $\delta_i^{(k)} < 0$, then rearrange (2.5) as,

$$|a_{ik}^{(k)}| = a_{ii}^{(k)} - \sum_{j=k+1, j \neq i}^n |a_{ij}^{(k)}| - v_i^{(k)}$$

and sum over i to obtain,

$$\begin{aligned} \sum_{i=k+1}^n |a_{ik}^{(k)}| &= \sum_{i=k+1}^n a_{ii}^{(k)} - \sum_{i=k+1}^n \sum_{j=k+1, j \neq i}^n |a_{ij}^{(k)}| - \sum_{i=k+1}^n v_i^{(k)} \\ &= \sum_{i=k+1}^n \sum_{j=k+1, j \neq i}^n |a_{ji}^{(k)}| + \sum_{i=k+1}^n \delta_i^{(k)} - \sum_{i=k+1}^n \sum_{j=k+1, j \neq i}^n |a_{ij}^{(k)}| - \sum_{i=k+1}^n v_i^{(k)} \\ &= \sum_{i=k+1}^n |a_{ki}^{(k)}| + \sum_{i=k+1}^n \delta_i^{(k)} - \sum_{i=k+1}^n v_i^{(k)} \\ &\leq a_{kk}^{(k)} + \sum_{i=k+1}^n \delta_i^{(k)} - \sum_{i=k+1}^n v_i^{(k)}. \end{aligned}$$

Rearranging yields

$$\begin{aligned}
\sum_{i=k+1}^n |a_{ik}^{(k)}| + v_i^{(k)} &\leq a_{kk}^{(k)} + \sum_{i=k+1, \delta_i^{(k)} \geq 0}^n \delta_i^{(k)} \leq a_{kk}^{(k)} + \sum_{i=k+1}^n a_{ii}^{(k)} \\
&\leq a_{kk}^{(k)} + \sum_{i=k+1}^n a_{kk}^{(k)} \leq a_{kk}^{(k)} + (n-k-1)a_{kk}^{(k)} \\
&= (n-k)a_{kk}^{(k)}
\end{aligned}$$

since for $\delta_i^{(k)} \geq 0, \delta_i^{(k)} \leq a_{ii}^{(k)} \leq a_{kk}^{(k)}$. □

To prove our improved perturbation bound for the LDU factorization, we need to consider perturbations that are more general in the p th column, for a fixed p . That is, we consider perturbations that satisfy

$$|\tilde{v} - v| \leq \epsilon v, |\tilde{a}_{ip} - a_{ip}| \leq \epsilon(v_i + |a_{ip}|), \text{ and } |\tilde{a}_{ij} - a_{ij}| \leq \epsilon|a_{ij}|. \quad (2.18)$$

for $i = 1, \dots, n$, $j = 1, \dots, n$, and $j \neq i, p$. This generalized perturbation can be equivalently expressed as

$$\tilde{v}_i = v_i(1 + \varphi_i) \text{ where } |\varphi_i| \leq \epsilon < 1 \text{ for } i = 1, \dots, n \quad (2.19)$$

$$\tilde{a}_{ip} = a_{ip}(1 + \varphi'_{ip}) + \varphi_{ip}v_i \text{ where } |\varphi_{ip}| \leq \epsilon < 1 \text{ and } \varphi'_{ip} = \varphi_{ip}\text{sign}(a_{ip}) \quad (2.20)$$

$$\tilde{a}_{ij} = a_{ij}(1 + \varphi_{ij}) \text{ where } |\varphi_{ij}| \leq \epsilon < 1 \text{ for } j \neq i, p \text{ and } j = 1, \dots, n \quad (2.21)$$

The following lemma is a generalization of Lemma 2.10

Lemma 2.15. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ satisfy (2.18). Suppose $\tilde{A}^{[i]} = \mathcal{D}(\tilde{A}_D^{[i]}, \tilde{v}^{[i]}) \in \mathbf{R}^{n \times n}$ is a matrix that differs from A in only the i th row and whose i th row is the same as the i th row of \tilde{A} . Then*

$$\det \tilde{A}^{[i]} = (\det A)(1 + \eta_i), \text{ where } |\eta_i| \leq 3\epsilon. \quad (2.22)$$

Proof. We consider the cofactor expansion of $\det \tilde{A}^{[i]}$ across row i . Let C_{ij} be the algebraic cofactors of $\tilde{A}^{[i]}$ corresponding to $\tilde{a}_{ij}^{[i]}$ and observe that they are also the algebraic cofactors of A corresponding to entry a_{ij} .

For $i = p$, use Lemma 2.9, (2.19), and (2.21) to show

$$\det \tilde{A}^{[p]} = \det A + \varphi_p v_p C_{pp} + \sum_{j \neq p} \varphi_{pj} (|a_{pj}| C_{pp} + a_{pj} C_{pj}).$$

Rearranging and applying the absolute value yields

$$\begin{aligned} |\det \tilde{A}^{[p]} - \det A| &= \left| \varphi_p v_p C_{pp} + \sum_{j \neq p} \varphi_{pj} (|a_{pj}| C_{pp} + a_{pj} C_{pj}) \right| \\ &\leq |\varphi_p| v_p C_{pp} + \sum_{j \neq p} |\varphi_{pj}| (|a_{pj}| C_{pp} + a_{pj} C_{pj}) \\ &\leq \epsilon v_p C_{pp} + \sum_{j \neq p} \epsilon (|a_{pj}| C_{pp} + a_{pj} C_{pj}) \end{aligned}$$

and hence $|\det \tilde{A}^{[p]} - \det A| \leq \epsilon \det A$

For $i \neq p$, use Lemma 2.9, (2.19), and (2.21) to obtain

$$\begin{aligned} \det \tilde{A}^{[i]} &= \tilde{v}_i C_{ii} + \sum_{j \neq i} |\tilde{a}_{ij}| C_{ii} + \tilde{a}_{ij} C_{ij} \\ &= \tilde{v}_i C_{ii} + |\tilde{a}_{ip}| C_{ii} + \tilde{a}_{ip} C_{ip} + \sum_{j \neq i, p} |\tilde{a}_{ij}| C_{ii} + \tilde{a}_{ij} C_{ij} \\ &= v_i C_{ii} + v_i \varphi_i C_{ii} + |\tilde{a}_{ip}| C_{ii} + a_{ip} C_{ip} + a_{ip} \varphi'_{ip} C_{ip} + \varphi_{ip} v_i C_{ip} \\ &\quad + \sum_{j \neq i, p} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) + \sum_{j \neq i, p} \varphi_{ij} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) \end{aligned}$$

From (2.20) we have

$$|\tilde{a}_{ip}| \leq |a_{ip}(1 + \varphi'_{ip})| + |\varphi_{ip}| v_i \leq |a_{ip}|(1 + \varphi'_{ip}) + \epsilon v_i$$

and

$$|\tilde{a}_{ip}| \geq |a_{ip}(1 + \varphi'_{ip})| - |\varphi_{ip} v_i| = |a_{ip}|(1 + \varphi'_{ip}) - |\varphi_{ip}| v_i \geq |a_{ip}|(1 + \varphi'_{ip}) - \epsilon v_i$$

That is,

$$|a_{ip}|(1 + \varphi'_{ip}) - \epsilon v_i \leq |\tilde{a}_{ip}| \leq |a_{ip}|(1 + \varphi'_{ip}) + \epsilon v_i \quad (2.23)$$

and hence

$$\begin{aligned}
\det \tilde{A}^{[i]} &\geq v_i C_{ii} + v_i \varphi_i C_{ii} + |a_{ip}| C_{ii} + |a_{ip}| \varphi'_{ip} C_{ii} - \epsilon v_i C_{ii} + a_{ip} C_{ip} + a_{ip} \varphi'_{ip} C_{ip} \\
&\quad + \varphi_{ip} v_i C_{ip} + \sum_{j \neq i, p} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) + \sum_{j \neq i, p} \varphi_{ij} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) \\
&= \det A + v_i \varphi_i C_{ii} + |a_{ip}| \varphi'_{ip} C_{ii} - \epsilon v_i C_{ii} + a_{ip} \varphi'_{ip} C_{ip} + \varphi_{ip} v_i C_{ip} \\
&\quad + \sum_{j \neq i, p} \varphi_{ij} (|a_{ij}| C_{ii} + a_{ij} C_{ij}).
\end{aligned}$$

Using Theorem 1.7, we have

$$\begin{aligned}
\det \tilde{A}^{[i]} &\geq \det A - \epsilon v_i C_{ii} - \epsilon |a_{ip}| C_{ii} - \epsilon v_i C_{ii} - \epsilon a_{ip} C_{ip} - \epsilon v_i C_{ii} \\
&\quad - \epsilon \sum_{j \neq i, p} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) \\
&\geq \det A - 3\epsilon v_i C_{ii} - \epsilon \sum_{j \neq i} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) \\
&\geq \det A - 3\epsilon \det A = \det A(1 - 3\epsilon)
\end{aligned}$$

and similarly,

$$\begin{aligned}
\det \tilde{A}^{[i]} &\leq v_i C_{ii} + v_i \varphi_i C_{ii} + |a_{ip}| C_{ii} + |a_{ip}| \varphi'_{ip} C_{ii} + \epsilon v_i C_{ii} + a_{ip} C_{ip} + a_{ip} \varphi'_{ip} C_{ip} \\
&\quad + \varphi_{ip} v_i C_{ip} + \sum_{j \neq i, p} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) + \sum_{j \neq i, p} \varphi_{ij} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) \\
&\leq v_i C_{ii} + \epsilon v_i C_{ii} + |a_{ip}| C_{ii} + \epsilon |a_{ip}| C_{ii} + \epsilon v_i C_{ii} + a_{ip} C_{ip} + \epsilon a_{ip} C_{ip} + \epsilon v_i C_{ip} \\
&\quad + \sum_{j \neq i, p} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) + \epsilon \sum_{j \neq i, p} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) \\
&\leq \det A + 3\epsilon v_i C_{ii} + \epsilon \sum_{j \neq i} (|a_{ij}| C_{ii} + a_{ij} C_{ij}) \\
&\leq \det A + 3\epsilon \det A = \det A(1 + 3\epsilon).
\end{aligned}$$

Thus,

$$-3\epsilon \det A \leq \det \tilde{A}^{[i]} - \det A \leq 3\epsilon \det A$$

or

$$|\det \tilde{A}^{[i]} - \det A| \leq 3\epsilon \det A$$

If $\det A = 0$ then $|\det \tilde{A}^{[i]} - \det A| = 0$ which implies $\det \tilde{A}^{[i]} = \det A = 0$. Thus, $\det \tilde{A}^{[i]} \leq \det A(1 + \eta_i)$ for any η_i and in particular, for $|\eta_i| \leq 3\epsilon$.

If $\det A > 0$ then

$$\det \tilde{A}^{[i]} = \det A(1 + \eta_i) = \det A + (\det A)\eta_i$$

which implies

$$\det \tilde{A}^{[i]} - \det A = (\det A)\eta_i$$

or

$$\eta_i = \frac{\det \tilde{A}^{[i]} - \det A}{\det A}.$$

Hence,

$$|\eta_i| = \frac{|\det \tilde{A}^{[i]} - \det A|}{\det A} \leq 3\epsilon$$

□

Lemma 2.16 uses Lemma 2.15 to present a similar perturbation bound for the principal minors of a diagonally dominant matrix with nonnegative diagonals under structured perturbations of type (2.18). As shown in [13, Lemma 4], these results hold for $|\eta_k^{(\alpha)}| \leq \epsilon$ for perturbation structures of type (2.7).

Lemma 2.16. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ satisfy (2.18). Suppose $\tilde{A}^{[i]} = \mathcal{D}(\tilde{A}_D^{[i]}, \tilde{v}^{[i]}) \in \mathbf{R}^{n \times n}$ is a matrix that differs from A in only the i th row and whose i th row is the same as the i th row of \tilde{A} . Let $1 \leq i_1 < i_2 < \dots < i_q \leq n$ and $\alpha = \{i_1, i_2, \dots, i_q\}$, and denote the principle submatrix of A that lies in rows and columns indexed by α as $A(\alpha, \alpha)$. Then we have*

$$\det \tilde{A}^{[i]}(\alpha, \alpha) = \begin{cases} \det A(\alpha, \alpha) & \text{if } i \notin \alpha \\ (\det A(\alpha, \alpha))(1 + \delta_i^{(\alpha)}) & \text{if } i \in \alpha \end{cases} \quad (2.24)$$

where $|\delta_i^{(\alpha)}| \leq 6\epsilon$ if $p \notin \alpha$ and $|\delta_i^{(\alpha)}| \leq 3\epsilon$ if $p \in \alpha$, and

$$\det \tilde{A}(\alpha, \alpha) = (\det A(\alpha, \alpha))(1 + \eta_1^{(\alpha)}) \cdots (1 + \eta_q^{(\alpha)}) \quad (2.25)$$

where $|\eta_k^{(\alpha)}| \leq 6\epsilon$ if $p \notin \alpha$ and $|\eta_k^{(\alpha)}| \leq 3\epsilon$ if $p \in \alpha$ for $k = 1, \dots, q$.

Proof. Again, we follow the proof of [13, Lemma 4]. Assume $i \in \alpha$, otherwise result is trivial. Since A , $\tilde{A}^{[i]}$, and \tilde{A} are diagonally dominant with nonnegative diagonals, then so are $A(\alpha, \alpha)$, $\tilde{A}^{[i]}(\alpha, \alpha)$, and $\tilde{A}(\alpha, \alpha)$, see Theorem 1.7. Hence, we can parameterize them in terms of their diagonally dominant parts and off diagonal entries. Let

$$A(\alpha, \alpha) = \mathcal{D}(A_D(\alpha, \alpha), w), \quad \tilde{A}^{[i]}(\alpha, \alpha) = \mathcal{D}(\tilde{A}_D^{[i]}(\alpha, \alpha), \tilde{w}^{[i]}), \quad \text{and}$$

$$\tilde{A}(\alpha, \alpha) = \mathcal{D}(\tilde{A}_D(\alpha, \alpha), \tilde{w})$$

where $w = [w_j]$, $\tilde{w}^{[i]} = [\tilde{w}_j^{[i]}]$, $\tilde{w} = [\tilde{w}_j] \in \mathbf{R}^q$. Observe that

$$w_i = a_{ii} - \sum_{j \in \alpha \setminus \{i\}} |a_{ij}| = \left(v_i + \sum_{j \neq i} |a_{ij}| \right) - \sum_{j \in \alpha \setminus \{i\}} |a_{ij}| = v_i + \sum_{j \notin \alpha} |a_{ij}|$$

and similarly $\tilde{w}_i^{[i]} = \tilde{v}_i^{[i]} + \sum_{j \notin \alpha} |\tilde{a}_{ij}^{[i]}|$. Thus, we have

$$\tilde{w}_i^{[i]} - w_i = \tilde{v}_i^{[i]} - v_i + \sum_{j \notin \alpha} |\tilde{a}_{ij}^{[i]}| - |a_{ij}|$$

and

$$\begin{aligned} |\tilde{w}_i^{[i]} - w_i| &\leq \left| \tilde{v}_i^{[i]} - v_i + \sum_{j \notin \alpha} |\tilde{a}_{ij}^{[i]}| - |a_{ij}| \right| \leq |\tilde{v}_i^{[i]} - v_i| + \sum_{j \notin \alpha} \left| |\tilde{a}_{ij}^{[i]}| - |a_{ij}| \right| \\ &\leq |\tilde{v}_i^{[i]} - v_i| + \sum_{j \notin \alpha} \left| \tilde{a}_{ij}^{[i]} - a_{ij} \right| \end{aligned}$$

If $p \in \alpha$, then

$$|\tilde{w}_i^{[i]} - w_i| \leq \epsilon v_i + \sum_{j \notin \alpha} \epsilon |a_{ij}| = \epsilon (v_i + \sum_{j \notin \alpha} |a_{ij}|) = \epsilon w_i.$$

If $p \notin \alpha$, then

$$\begin{aligned} |\tilde{w}_i^{[i]} - w_i| &\leq |\tilde{v}_i^{[i]} - v_i| + \sum_{j \notin \alpha, j \neq p} \left| \tilde{a}_{ij}^{[i]} - a_{ij} \right| + |\tilde{a}_{ip}^{[i]} - a_{ip}| \\ &\leq \epsilon v_i + \sum_{j \notin \alpha, j \neq p} \epsilon |a_{ij}| + \epsilon (v_i + |a_{ip}|) \\ &= 2\epsilon v_i + \sum_{j \notin \alpha} \epsilon |a_{ij}| \leq 2\epsilon (v_i + \sum_{j \notin \alpha} |a_{ij}|) = 2\epsilon w_i \end{aligned}$$

Hence, if $p \in \alpha$, then $|\tilde{w}_k^{[i]} - w_k| \leq \epsilon w_k$ for all $k \in \alpha$. In addition, $w_k \geq v_k$ and so, the off-diagonal entries of $\tilde{A}^{[i]}(\alpha, \alpha)$ and $A(\alpha, \alpha)$ satisfy conditions (2.18) for their parameters. Therefore, we can apply Lemma 2.15 to $\tilde{A}^{[i]}(\alpha, \alpha)$ and $A(\alpha, \alpha)$ to obtain

$$\det \tilde{A}^{[i]}(\alpha, \alpha) = (\det A(\alpha, \alpha))(1 + \delta_i^{(\alpha)})$$

with $|\delta_i^{(\alpha)}| \leq 3\epsilon$.

On the other hand, if $p \notin \alpha$, then $|\tilde{w}_k^{[i]} - w_k| \leq 2\epsilon w_k$ for all $k \in \alpha$. Again, the off-diagonal entries of $\tilde{A}^{[i]}(\alpha, \alpha)$ and $A(\alpha, \alpha)$ satisfy (2.18) for their parameters. So, we can apply Lemma 2.15 to $\tilde{A}^{[i]}(\alpha, \alpha)$ and $A(\alpha, \alpha)$, but this time with ϵ replaced by 2ϵ , which requires $2\epsilon < 1$, to obtain

$$\det \tilde{A}^{[i]}(\alpha, \alpha) = (\det A(\alpha, \alpha))(1 + \delta_i^{(\alpha)})$$

with $|\delta_i^{(\alpha)}| \leq 6\epsilon$.

For $\det A(\alpha, \alpha)$, we use the same strategy as in [13, Lemma 3b]. Consider that the perturbed submatrix $\tilde{A}(\alpha, \alpha)$ can be obtained from $A(\alpha, \alpha)$ by a sequence of “only one row” at a time perturbations. By (2.24), each of these “only one row” perturbations produces a determinant that is equal to the determinant before the perturbation times a factor $1 + \eta$, with $|\eta| \leq 6\epsilon$ if $p \notin \alpha$ and $|\eta| \leq 3\epsilon$ if $p \in \alpha$. \square

Lemma 2.17. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $1 \leq k \leq n - 2$, $k + 1 \leq p, q \leq n$ and $p \neq q$. Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ that satisfies (2.18) with the given p . Suppose $\tilde{A}^{[i]} = \mathcal{D}(\tilde{A}_D^{[i]}, \tilde{v}^{[i]}) \in \mathbf{R}^{n \times n}$ is a matrix that differs from A in only the i th row and whose i th row parameters are the entries in the i th row of \tilde{A} . Then, the following statements hold*

$$1. \left| (\tilde{g}^{[i]})_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \leq \begin{cases} 0 & \text{if } i \notin \{1, \dots, k, p\} \\ 4\epsilon g_{pp}^{(k+1)} & \text{if } i \in \{1, \dots, k, p\} \end{cases}$$

$$2. |\tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)}| \leq \frac{4}{3} ((1 + 3\epsilon)^{k+1} - 1) g_{pp}^{(k+1)}.$$

Proof. Let G_{ij} be the algebraic cofactor of $A([1 : k, p], [1 : k, q])$ corresponding to entry a_{ij} and note that these are also the algebraic cofactors of $\tilde{A}^{[i]}([1 : k, p], [1 : k, q])$ corresponding to entry $\tilde{a}_{ij}^{[i]}$. For $1 \leq i \leq k$, applying (2.14) yields

$$\begin{aligned} (\tilde{g}^{[i]})_{pq}^{(k+1)} &= \left(\tilde{v}_i^{[i]} + \sum_{j \notin \{1, \dots, k, q\}} |\tilde{a}_{ij}^{[i]}| \right) G_{ii} + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} \tilde{a}_{ij}^{[i]} G_{ij} + |\tilde{a}_{ij}^{[i]}| G_{ii} \\ &= \left(\tilde{v}_i + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij}| \right) G_{ii} + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} (\tilde{a}_{ij} G_{ij} + |\tilde{a}_{ij}| G_{ii}) + |\tilde{a}_{ip}| G_{ii} \end{aligned}$$

and then use the discussion following (2.18) to obtain

$$\begin{aligned} (\tilde{g}^{[i]})_{pq}^{(k+1)} &= \left(v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \right) G_{ii} + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} (a_{ij} G_{ij} + |a_{ij}| G_{ii}) \\ &\quad + \left(v_i \varphi_i + \sum_{j \notin \{1, \dots, k, p, q\}} \varphi_{ij} |a_{ij}| \right) G_{ii} \\ &\quad + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} \varphi_{ij} (a_{ij} G_{ij} + |a_{ij}| G_{ii}) + |\tilde{a}_{ip}| G_{ii} \end{aligned} \tag{2.26}$$

We find an upper bound for $(\tilde{g}^{[i]})_{pq}^{(k+1)}$ for $1 \leq i \leq k$ by considering two cases. First, suppose $G_{ii} < 0$. We make use of a simple property: Let $b \leq c$ be real numbers and $G_{ii} < 0$. Then $cG_{ii} \leq bG_{ii}$. This means that if a lower bound of $|\tilde{a}_{ip}|$ replaces $|\tilde{a}_{ip}|$ in (2.26), then an upper bound for $(\tilde{g}^{[i]})_{pq}^{(k+1)}$ is obtained. In particular, if we use in

(2.26) the lower bound of $|\tilde{a}_{ip}|$ appearing in (2.23) then

$$\begin{aligned}
(\tilde{g}^{[i]})_{pq}^{(k+1)} &\leq g_{pq}^{(k+1)} + \left(v_i \varphi_i + \sum_{j \notin \{1, \dots, k, p, q\}} \varphi_{ij} |a_{ij}| \right) G_{ii} \\
&\quad + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} \varphi_{ij} (a_{ij} G_{ij} + |a_{ij}| G_{ii}) + |a_{ip}| \varphi'_{ip} G_{ii} - \epsilon v_i G_{ii} \\
&\leq g_{pq}^{(k+1)} + \epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \right) |G_{ii}| \right. \\
&\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| + |a_{ip}| |G_{ii}| + v_i |G_{ii}| \right] \\
&\leq g_{pq}^{(k+1)} + 2\epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right) |G_{ii}| \right. \\
&\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| \right].
\end{aligned}$$

Applying Lemma 2.12 yields

$$(\tilde{g}^{[i]})_{pq}^{(k+1)} \leq g_{pq}^{(k+1)} + 4\epsilon g_{pp}^{(k+1)} \quad (2.27)$$

Now suppose $G_{ii} \geq 0$ and substitute the upper bound for $|\tilde{a}_{ip}|$ in (2.23) into (2.26) to get

$$\begin{aligned}
(\tilde{g}^{[i]})_{pq}^{(k+1)} &\leq g_{pq}^{(k+1)} + \left(v_a \varphi_i + \sum_{j \notin \{1, \dots, k, p, q\}} \varphi_{ij} |a_{ij}| \right) G_{ii} \\
&\quad + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} \varphi_{ij} (a_{ij} G_{ij} + |a_{ij}| G_{ii}) + |a_{ip}| \varphi'_{ip} G_{ii} + \epsilon v_i G_{ii} \\
&\leq g_{pq}^{(k+1)} + \epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \right) |G_{ii}| \right. \\
&\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| + |a_{ip}| |G_{ii}| + v_i |G_{ii}| \right] \\
&\leq g_{pq}^{(k+1)} + 2\epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right) |G_{ii}| \right. \\
&\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| \right].
\end{aligned}$$

Again, applying Lemma 2.12 gives

$$(\tilde{g}^{[i]})_{pq}^{(k+1)} \leq g_{pq}^{(k+1)} + 4\epsilon g_{pp}^{(k+1)}. \quad (2.28)$$

We find a lower bound for $(\tilde{g}^{[i]})_{pq}^{(k+1)}$ by again considering two cases. Suppose $G_{ii} < 0$.

We use the upper bound in (2.23) for $|\tilde{a}_{ip}|$ to obtain a lower bound on $|\tilde{a}_{ip}|G_{ii}$

$$\begin{aligned} (\tilde{g}^{[i]})_{pq}^{(k+1)} &\geq g_{pq}^{(k+1)} + \left(v_i \varphi_i + \sum_{j \notin \{1, \dots, k, p, q\}} \varphi_{ij} |a_{ij}| \right) G_{ii} \\ &\quad + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} \varphi_{ij} (a_{ij} G_{ij} + |a_{ij}| G_{ii}) + |a_{ip}| \varphi'_{ip} G_{ii} + \epsilon v_i G_{ii} \\ &\geq g_{pq}^{(k+1)} - \epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \right) |G_{ii}| \right. \\ &\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| + |a_{ip}| |G_{ii}| + v_i |G_{ii}| \right] \\ &\geq g_{pq}^{(k+1)} - 2\epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right) |G_{ii}| \right. \\ &\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| \right]. \end{aligned}$$

Applying Lemma 2.12 gives

$$(\tilde{g}^{[i]})_{pq}^{(k+1)} \geq g_{pq}^{(k+1)} - 4\epsilon g_{pp}^{(k+1)}. \quad (2.29)$$

Now suppose $G_{ii} \geq 0$ and use the lower bound in (2.23) for $|\tilde{a}_{ip}|$ to get a lower bound

for $|\tilde{a}_{ip}|G_{ii}$:

$$\begin{aligned}
(\tilde{g}^{[i]})_{pq}^{(k+1)} &\geq g_{pq}^{(k+1)} + \left(v_i \varphi_i + \sum_{j \notin \{1, \dots, k, p, q\}} \varphi_{ij} |a_{ij}| \right) G_{ii} \\
&\quad + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} \varphi_{ij} (a_{ij} G_{ij} + |a_{ij}| G_{ii}) + |a_{ip}| \varphi'_{ip} G_{ii} - \epsilon v_i G_{ii} \\
&\geq g_{pq}^{(k+1)} - \epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \right) |G_{ii}| \right. \\
&\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| + |a_{ip}| |G_{ii}| + v_i |G_{ii}| \right] \\
&\geq g_{pq}^{(k+1)} - 2\epsilon \left[\left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right) |G_{ii}| \right. \\
&\quad \left. + \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij} G_{ij} + |a_{ij}| G_{ii}| \right].
\end{aligned}$$

Apply Lemma 2.12 to obtain

$$(\tilde{g}^{[i]})_{pq}^{(k+1)} \geq g_{pq}^{(k+1)} - 4\epsilon g_{pp}^{(k+1)}. \quad (2.30)$$

Hence, by (2.27), (2.28), (2.29) and (2.30) we have for $1 \leq i \leq k$

$$g_{pq}^{(k+1)} - 4\epsilon g_{pp}^{(k+1)} \leq (\tilde{g}^{[i]})_{pq}^{(k+1)} \leq g_{pq}^{(k+1)} + 4\epsilon g_{pp}^{(k+1)}$$

or,

$$-4\epsilon g_{pp}^{(k+1)} \leq (\tilde{g}^{[i]})_{pq}^{(k+1)} - g_{pq}^{(k+1)} \leq 4\epsilon g_{pp}^{(k+1)}$$

which gives

$$|(\tilde{g}^{[i]})_{pq}^{(k+1)} - g_{pq}^{(k+1)}| \leq 4\epsilon g_{pp}^{(k+1)}.$$

Now, for $i = p$, use (2.12) to obtain

$$\begin{aligned}
(\tilde{g}^{[p]})_{pq}^{(k+1)} &= \sum_{j \in \{1, \dots, k, q\}} \tilde{a}_{pj}^{[p]} G_{pj} = \sum_{j \in \{1, \dots, k, q\}} a_{pj} (1 + \varphi_{pj}) G_{pj} \\
&= \sum_{j \in \{1, \dots, k, q\}} a_{pj} G_{pj} + \sum_{j \in \{1, \dots, k, q\}} \varphi_{pj} a_{pj} G_{pj} = g_{pq}^{(k+1)} + \sum_{j \in \{1, \dots, k, q\}} \varphi_{pj} a_{pj} G_{pj}
\end{aligned}$$

Thus,

$$\left| \left(\tilde{g}^{[p]} \right)_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \leq \epsilon \sum_{j \in \{1, \dots, k, q\}} |a_{pj} G_{pj}|. \quad (2.31)$$

Applying (2.13) gives

$$\left| \left(\tilde{g}^{[p]} \right)_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \leq \epsilon (2g_{pp}^{(k+1)}) \quad (2.32)$$

and thus

$$\left| \left(\tilde{g}^{[i]} \right)_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \leq 4\epsilon (g_{pp}^{(k+1)}) \quad (2.33)$$

for all $i \in [1 : k, p]$.

Part (2). Consider obtaining \tilde{A} from A by a sequence of only one row at a time perturbations. Note that each matrix in this sequence is row diagonally dominant with nonnegative diagonals. The variation in $g_{pq}^{(k+1)}$ is a consequence only of the perturbations of rows with indices in $\{1, \dots, k, p\}$. Let α be a subset of $\{1, \dots, k, p\}$ and denote by $(\tilde{g}^\alpha)_{pq}^{(k+1)}$ the minor corresponding to a matrix obtained from A through perturbations in the rows with indices in α only.

$$\begin{aligned} \left| \tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| &= \left| \left(\tilde{g}^{\{1, \dots, k, p\}} \right)_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \\ &\leq \left| \left(\tilde{g}^{\{1, \dots, k, p\}} \right)_{pq}^{(k+1)} - \left(\tilde{g}^{\{1, \dots, k\}} \right)_{pq}^{(k+1)} \right| + \dots + \left| \left(\tilde{g}^{\{1\}} \right)_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right|. \end{aligned}$$

Apply Lemma 2.17 to each term in the sum to obtain,

$$\left| \tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \leq 4\epsilon \left[\left(\tilde{g}^{\{1, \dots, k\}} \right)_{pp}^{(k+1)} + \dots + g_{pp}^{(k+1)} \right]$$

and then apply Lemma 2.16 to each term in the sum to obtain

$$\left| \tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)} \right| \leq 4\epsilon \left[(1 + 3\epsilon)^k + \dots + 1 \right] g_{pp}^{(k+1)} \leq \frac{4}{3} \left((1 + 3\epsilon)^{k+1} - 1 \right) g_{pp}^{(k+1)} \quad (2.34)$$

□

In Theorem 2.20, we present a perturbation bound for the minors of a diagonally dominant matrix with nonnegative diagonals with structure perturbation of type (2.7). Our proof depends upon the construction of a new diagonally dominant matrix with nonnegative diagonals with structured perturbation of type (2.18).

Lemma 2.18. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ be a matrix that satisfies*

$$|\tilde{v} - v| \leq \epsilon v \text{ and } |\tilde{A}_D - A_D| \leq \epsilon |A_D|, \text{ for some } 0 \leq \epsilon < \frac{1}{3}. \quad (2.35)$$

Define $B = [b_{ij}] \in \mathbf{R}^{(k+2) \times (k+2)}$ by

$$b_{ij} = \begin{cases} a_{ij} & \text{for } i \in \{1, \dots, k, p, q\} \text{ and } j \in \{1, \dots, k, q\} \\ a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} & \text{for } i \in \{1, \dots, k, p, q\} \text{ and } j = p \end{cases} \quad (2.36)$$

and define $\tilde{B} = [\tilde{b}_{ij}] \in \mathbf{R}^{(k+2) \times (k+2)}$ by

$$\tilde{b}_{ij} = \begin{cases} \tilde{a}_{ij} & \text{for } i \in \{1, \dots, k, p, q\} \text{ and } j \in \{1, \dots, k, q\} \\ \tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} & \text{for } i \in \{1, \dots, k, p, q\} \text{ and } j = p \end{cases} \quad (2.37)$$

where $s_j = \text{sign}(a_{pj}^{(k+1)})$. Then B and \tilde{B} are diagonally dominant matrices that can be parameterized as $B = \mathcal{D}(B_D, w)$ and $\tilde{B} = \mathcal{D}(\tilde{B}_D, \tilde{w})$ that satisfy

$$|\tilde{w} - w| \leq \delta w, |\tilde{b}_{ip} - b_{ip}| \leq \delta(w_i + |b_{ip}|), \text{ and } |\tilde{b}_{ij} - b_{ij}| \leq \delta |b_{ij}|$$

for $i \in \{1, \dots, k, p, q\}$, $j \in \{1, \dots, k, q\}$, and $j \neq i$ and $\delta = \frac{2\epsilon}{1-\epsilon}$.

Proof. Note that we label the $k+2$ rows and columns of B using the indices $\{1, \dots, k, p, q\}$. While not traditional, this labeling is useful because we can easily compare entries in B to entries in A . Note that the $(k+1)$ th row and column of B correspond to the p th row and column of A and similarly the $(k+2)$ nd row and column of B is the q th row and column of A .

Using the diagonal dominance of A , for $1 \leq i \leq k$ we have

$$\begin{aligned}
\sum_{j \in \{1, \dots, k, p, q\} \setminus \{i\}} |b_{ij}| &= \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |b_{ij}| + |b_{ip}| \\
&= \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij}| + |a_{ip}| - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \\
&\leq \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij}| + |a_{ip}| + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \\
&= \sum_{j \neq i} |a_{ij}| \leq a_{ii} = b_{ii}
\end{aligned}$$

and for $i = p$

$$\begin{aligned}
\sum_{j \in \{1, \dots, k, q\}} |b_{pj}| &= \sum_{j \in \{1, \dots, k, q\}} |a_{pj}| = \sum_{j \neq p} |a_{pj}| - \sum_{j \notin \{1, \dots, k, p, q\}} |a_{pj}| \\
&\leq \sum_{j \neq p} |a_{pj}| - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{pj} \leq a_{pp} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{pj} = b_{pp}
\end{aligned}$$

Hence, B is diagonally dominant with nonnegative diagonals. Using the same argument and the diagonal dominance of \tilde{A} , we can show \tilde{B} is diagonally dominant with nonnegative diagonals as well. Thus, we can parameterize B and \tilde{B} in terms of their diagonally dominant parts and off diagonal entries. Let

$$B = \mathcal{D}(B_D, w) \text{ and } \tilde{B} = \mathcal{D}(\tilde{B}_D, \tilde{w})$$

where $w, \tilde{w} \in \mathbf{R}^{k+2}$. Now, note for $i \in \{1, \dots, k, p, q\}$, $j \in \{1, \dots, k, q\}$, and $i \neq j$ we have

$$|\tilde{b}_{ij} - b_{ij}| = |\tilde{a}_{ij} - a_{ij}| \leq \epsilon |a_{ij}| = \epsilon |b_{ij}|$$

and for $i \in \{1, \dots, k, p, q\}$, $j = p$, and $i \neq j$

$$\begin{aligned}
|\tilde{b}_{ip} - b_{ip}| &= \left| \left(\tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \right) - \left(a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right) \right| \\
&\leq |\tilde{a}_{ip} - a_{ip}| + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij} - a_{ij}| \leq \epsilon (|a_{ip}| + v_i) + \epsilon \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \\
&\leq \epsilon \left(|a_{ip}| + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| + v_i \right) \\
&= \epsilon \left(\sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| + v_i \right) = \epsilon \left(\sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| + a_{ii} - \sum_{j \neq p} |a_{ij}| \right) \\
&= \epsilon \left(a_{ii} - \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij}| \right) = \epsilon \left(b_{ii} - \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |b_{ij}| \right) \\
&= \epsilon \left(b_{ii} - \sum_{j \in \{1, \dots, k, p, q\} \setminus \{i\}} |b_{ij}| + |b_{ip}| \right) = \epsilon (w_i + |b_{ip}|).
\end{aligned}$$

Thus, the off diagonal entries of \tilde{B} are perturbations of the off diagonal entries of B of type (2.18). Now we focus on the diagonally dominant parts. Let $i \in \{1, \dots, k, q\}$ and observe

$$\begin{aligned}
w_i &= b_{ii} - \sum_{j \in \{1, \dots, k, p, q\} \setminus \{i\}} |b_{ij}| = b_{ii} - \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |b_{ij}| - |b_{ip}| \\
&= a_{ii} - \sum_{j \in \{1, \dots, k, q\} \setminus \{i\}} |a_{ij}| - \left| a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right| \\
&= v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| - \left| a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right| \\
&= v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| + |a_{ip}| - \left| a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right|
\end{aligned}$$

Similarly, we have

$$\tilde{w}_i = \tilde{v}_i + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij}| + |\tilde{a}_{ip}| - |\tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij}|.$$

Let $i \in \{1, \dots, k, q\}$. We will consider two cases.

Case 1: $\text{sign} \left(a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right) = \text{sign} \left(\tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \right) := \theta$. Then,

$$\begin{aligned}
\tilde{w}_i &= \tilde{v}_i + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij}| + |\tilde{a}_{ip}| - \theta (\tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij}) \\
&= \tilde{v}_i + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij}| + |\tilde{a}_{ip}| - \theta \tilde{a}_{ip} + \theta \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \\
&= \tilde{v}_i + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij}| (1 + \theta s_j \text{sign}(\tilde{a}_{ij})) + |\tilde{a}_{ip}| (1 - \theta \text{sign}(\tilde{a}_{ip})) \\
&= \tilde{v}_i + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij}| (1 + \theta s_j \text{sign}(a_{ij})) + |\tilde{a}_{ip}| (1 - \theta \text{sign}(a_{ip})).
\end{aligned}$$

since $\text{sign}(\tilde{a}_{ij}) = \text{sign}(a_{ij})$ for all $j \neq i$. Similarly, we have

$$w_i = v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| (1 + \theta s_j \text{sign}(a_{ij})) + |a_{ip}| (1 - \theta \text{sign}(a_{ip}))$$

and hence

$$\begin{aligned}
|\tilde{w}_i - w_i| &\leq |\tilde{v}_i - v_i| + \sum_{j \notin \{1, \dots, k, p, q\}} \left| |\tilde{a}_{ij}| - |a_{ij}| \right| (1 + \theta s_j \text{sign}(a_{ij})) \\
&\quad + \left| |\tilde{a}_{ip}| - |a_{ip}| \right| (1 - \theta \text{sign}(a_{ip})) \\
&\leq |\tilde{v}_i - v_i| + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij} - a_{ij}| (1 + \theta s_j \text{sign}(a_{ij})) \\
&\quad + |\tilde{a}_{ip} - a_{ip}| (1 - \theta \text{sign}(a_{ip})) \\
&\leq \epsilon v_i + \epsilon \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| (1 + \theta s_j \text{sign}(a_{ij})) + \epsilon |a_{ip}| (1 - \theta \text{sign}(a_{ip})) \\
&\leq \epsilon \left(v_i + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| + |a_{ip}| - \theta a_{ip} + \theta \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right) \\
&\leq \epsilon w_i
\end{aligned}$$

Case 2: $\text{sign}(a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij}) \neq \text{sign}(\tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij})$. Note

$$\begin{aligned}
& \left| \tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \right| + \left| a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right| \\
&= \left| \left(\tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \right) - \left(a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right) \right| \\
&\leq |\tilde{a}_{ip} - a_{ip}| + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{ij} - a_{ij}| \\
&\leq \epsilon |a_{ip}| + \epsilon \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}|
\end{aligned}$$

That is,

$$\left| \tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \right| + \left| a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right| \leq \epsilon \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \quad (2.38)$$

which yields

$$\begin{aligned}
|\tilde{w}_i - w_i| &= \left| \left(\tilde{v}_i + \sum_{j \notin \{1, \dots, k, q\}} |\tilde{a}_{ij}| - \left| \tilde{a}_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \right| \right) \right. \\
&\quad \left. - \left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| - \left| a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right| \right) \right| \\
&\leq |\tilde{v}_i - v_i| + \sum_{j \notin \{1, \dots, k, q\}} ||\tilde{a}_{ij}| - |a_{ij}|| \\
&\quad + \left| \left(\tilde{a}_{ip} + \sum_{j \notin \{1, \dots, k, p, q\}} s_j \tilde{a}_{ij} \right) - \left(a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right) \right| \\
&\leq \epsilon v_i + \epsilon \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| + \epsilon \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \\
&\leq 2\epsilon \left(v_i \sum_{j \notin \{1, \dots, k, p, q\}} |a_{ij}| \right)
\end{aligned}$$

From (2.38) we also have

$$\left| a_{ip} + \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} \right| \leq \epsilon \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}|$$

So,

$$\begin{aligned}
w_i &= v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| - |a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij}| \\
&\geq v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| - \epsilon \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \\
&\geq (1 - \epsilon) \left(v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right).
\end{aligned}$$

Rearranging gives

$$v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \leq \frac{w_i}{1 - \epsilon}$$

and thus we have

$$|\tilde{w}_i - w_i| \leq 2\epsilon \left[v_i + \sum_{j \notin \{1, \dots, k, q\}} |a_{ij}| \right] \leq \frac{2\epsilon}{1 - \epsilon} w_i$$

for $i \in \{1, \dots, k, q\}$. Now, we consider $i = p$. Note that

$$\begin{aligned}
w_p &= b_{pp} - \sum_{j \in \{1, \dots, k, q\}} |b_{pj}| = a_{pp} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{pj} - \sum_{j \in \{1, \dots, k, q\}} |a_{pj}| \\
&= v_p + \sum_{j \neq p} |a_{pj}| - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{pj} - \sum_{j \in \{1, \dots, k, q\}} |a_{pj}| \\
&= v_p + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{pj}| - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{pj} = v_p + \sum_{j \notin \{1, \dots, k, p, q\}} (|a_{pj}| - s_j a_{pj}) \\
&= v_p + \sum_{j \notin \{1, \dots, k, p, q\}} |a_{pj}| (1 - s_j \text{sign}(a_{pj}))
\end{aligned}$$

Similarly, we have

$$\tilde{w}_p = \tilde{v}_p + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{pj}| (1 - s_j \text{sign}(\tilde{a}_{pj})) = \tilde{v}_p + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{pj}| (1 - s_j \text{sign}(a_{pj}))$$

since $\text{sign}(\tilde{a}_{pj}) = \text{sign}(a_{pj})$. Thus,

$$\begin{aligned}
|\tilde{w}_p - w_p| &\leq |\tilde{v}_p - v_p| + \sum_{j \notin \{1, \dots, k, p, q\}} |\tilde{a}_{pj} - a_{pj}| (1 - s_j \text{sign}(a_{pj})) \\
&\leq \epsilon v_i + \epsilon \sum_{j \notin \{1, \dots, k, p, q\}} |a_{pj}| (1 - s_j \text{sign}(a_{pj})) = \epsilon w_p
\end{aligned}$$

So, we have that $|\tilde{w}_i - w_i| \leq \frac{2\epsilon}{1-\epsilon} w_i$ for all $i \in \{1, \dots, k, p, q\}$. Hence, \tilde{B} is a perturbation of B that satisfies

$$|\tilde{w} - w| \leq \delta w, \quad |\tilde{b}_{ip} - b_{ip}| \leq \delta(w_i + |b_{ip}|), \quad \text{and} \quad |\tilde{b}_{ij} - b_{ij}| \leq \delta |b_{ij}| \quad (2.39)$$

for $i \in \{1, \dots, k, p, q\}$ and $j \in [1 : k, q]$ where $\delta = \frac{2\epsilon}{1-\epsilon}$. Note that $\delta \geq 0$ since $0 \leq \epsilon < 1$ and if $\epsilon < \frac{1}{3}$ then $\delta < 1$. \square

Lemma 2.19. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $B = [b_{ij}] \in \mathbf{R}^{(k+2) \times (k+2)}$ be defined as*

$$b_{ij} = \begin{cases} a_{ij} & \text{for } i \in \{1, \dots, k, p, q\} \text{ and } j \in \{1, \dots, k, q\} \\ a_{ip} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{ij} & \text{for } i \in \{1, \dots, k, p, q\} \text{ and } j = p \end{cases}$$

where $s_j = \text{sign}(a_{pj}^{(k+1)})$ and row and column indices of B are indexed as $\{1, \dots, k, p, q\}$ with $k+1 \leq p, q \leq n$ and $p \neq q$. Let us define

$$(g_B)_{pp}^{(k+1)} := \det B([1 : k, p], [1 : k, p])$$

Then:

1. If $g_{kk}^{(k)} \neq 0$, then

$$(g_B)_{pp}^{(k+1)} = (v_p^{(k+1)} + |a_{pq}^{(k+1)}|) g_{kk}^{(k)}$$

2. If $g_{kk}^{(k)} = 0$ then $(g_B)_{pp}^{(k+1)} = 0$

Proof. Observe that $B([1 : k, p], [1 : k, p])$ and $A([1 : k, p], [1 : k, p])$ have columns 1 through k equal. In fact,

$$B([1 : k, p], [1 : k, p]) = A([1 : k, p], 1 : k), A([1 : k, p], p) - \sum_{j \notin \{1, \dots, k, p, q\}} s_j A([1 : k, p], j).$$

Use that the determinant is a linear function of any of its columns, assuming that the remaining columns are fixed, to obtain

$$(g_B)_{pp}^{(k+1)} = \det A([1 : k, p], [1 : k, p]) - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \det A([1 : k, p], [1 : k, j]) \quad (2.40)$$

If $g_{kk}^{(k)} := \det A(1 : k, 1 : k) \neq 0$, then

$$(g_B)_{pp}^{(k+1)} = g_{kk}^{(k)} \left(\frac{\det A([1 : k, p], [1 : k, p])}{\det A(1 : k, 1 : k)} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j \frac{\det A([1 : k, p], [1 : k, j])}{\det A(1 : k, 1 : k)} \right)$$

By (2.6)

$$\begin{aligned} (g_B)_{pp}^{(k+1)} &= g_{kk}^{(k)} \left(a_{pp}^{(k+1)} - \sum_{j \notin \{1, \dots, k, p, q\}} s_j a_{pj}^{(k+1)} \right) \\ &= g_{kk}^{(k)} \left(a_{pp}^{(k+1)} - \sum_{j \notin \{1, \dots, k, p, q\}} |a_{pj}^{(k+1)}| \right) \end{aligned}$$

Since, $a_{pj}^{(k+1)} = 0$ for $1 \leq j \leq k$ then

$$\begin{aligned} (g_B)_{kk}^{(k)} &= g_{kk}^{(k)} \left(a_{pp}^{(k+1)} - \sum_{j \neq p} |a_{pj}^{(k+1)}| + |a_{pq}^{(k+1)}| \right) \\ &= g_{kk}^{(k)} (v_p^{(k+1)} + |a_{pq}^{(k+1)}|) \end{aligned}$$

If $g_{kk}^{(k)} = 0$, then

$$\det A([1 : k, p], [1 : k, p]) = 0. \quad (2.41)$$

To see this, use that $A([1 : k, p], [1 : k, p])$ is row diagonally dominant with nonnegative diagonals, expand $\det A([1 : k, p], [1 : k, p])$ by cofactors across the last row and use Theorem 1.7 which gives

$$0 = |\det A(1 : k, 1 : k)| \geq |\det A(1 : k, [1 : j - 1, j + 1 : k, p])|$$

for $j = 1, \dots, k$.

In addition,

$$\det A([1 : k, p], [1 : k, j]) = 0 \quad (2.42)$$

for $j \notin \{1, \dots, k, p, q\}$. To see this, apply Theorem 1.7 to the row diagonally dominant matrix $A([1 : k, p, j], [1 : k, p, j])$ and get

$$0 = |\det A([1 : k, p], [1 : k, p])| \geq |\det A([1 : k, p], [1 : k, j])|$$

Combine (2.41) and (2.42) with (2.40) to obtain $(g_B)_{pp}^{(k+1)} = 0$. \square

Theorem 2.20. Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ be a matrix that satisfies

$$|\tilde{v} - v| \leq \epsilon v \text{ and } |\tilde{A}_D - A_D| \leq \epsilon |A_D| \quad (2.43)$$

for some ϵ with $0 \leq \epsilon < \frac{1}{3}$. Let $1 \leq k \leq n-2$, $k+1 \leq p, q \leq n$ and $p \neq q$, then if $g_{kk}^{(k)} \neq 0$

$$|\tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)}| \leq \frac{4}{3} ((1 + \epsilon_0)^{k+1} - 1) (v_p^{(k+1)} + |a_{pq}^{(k+1)}|) g_{kk}^{(k)} \quad (2.44)$$

where $\epsilon_0 = \frac{6\epsilon}{1-\epsilon}$. If $g_{kk}^{(k)} = 0$, then

$$\tilde{g}_{pq}^{(k+1)} = g_{pq}^{(k+1)} = 0 \quad (2.45)$$

Proof. Suppose $g_{pq}^{(k+1)} \neq 0$. Define B and \tilde{B} as in Lemma 2.18. By the results of 2.18 we can apply Lemma 2.17 to obtain

$$|(\tilde{g}_B)_{pq}^{(k+1)} - (g_B)_{pq}^{(k+1)}| \leq \frac{4}{3} ((1 + 3\delta)^{k+1} - 1) (g_B)_{pp}^{(k+1)}$$

By the construction of B and \tilde{B} , we have

$$(\tilde{g}_B)_{pq}^{(k+1)} = \tilde{g}_{pq}^{(k+1)} \text{ and } (g_B)_{pq}^{(k+1)} = g_{pq}^{(k+1)}$$

where $\delta = \frac{2\epsilon}{1-\epsilon}$ and hence

$$|\tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)}| \leq \frac{4}{3} ((1 + 3\delta)^{k+1} - 1) (g_B)_{pp}^{(k+1)}. \quad (2.46)$$

Apply Lemma 2.19 to obtain

$$|\tilde{g}_{pq}^{(k+1)} - g_{pq}^{(k+1)}| \leq \frac{4}{3} ((1 + 3\delta)^{k+1} - 1) (v_p^{(k+1)} + |a_{pq}^{(k+1)}|) g_{kk}^{(k)}.$$

We have seen in the proof of Lemma 2.19 that $g_{kk}^{(k)} = 0$ implies $\det A([1 : k, p], [1 : k, p]) = 0$. That is, $g_{pp}^{(k+1)} = 0$ and thus by Lemma 2.12, $g_{pq}^{(k+1)} = 0$ and by Lemma 2.17, we have $\tilde{g}_{pq}^{(k+1)} = g_{pq}^{(k+1)} = 0$. \square

Theorem 2.21. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$. Suppose A has LDU factorization $A = LDU$ and assume A is arranged for column diagonal dominance pivoting, i.e.*

$$a_{kk}^{(k)} = \max\{a_{ii}^{(k)} : a_{ii}^{(k)} - \sum_{j=k, j \neq i}^n |a_{ji}^{(k)}| \geq 0, k \leq i \leq n\}$$

Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ be a matrix that satisfies

$$|\tilde{A}_D - A_D| \leq \epsilon |A_D| \text{ and } |\tilde{v} - v| \leq \epsilon v$$

for some ϵ with $0 \leq \epsilon < \frac{1}{3}$. Assume $2n\epsilon_0 \leq 1$, where $\epsilon_0 = \frac{6\epsilon}{1-\epsilon}$. Then, we have

1. \tilde{A} is row diagonally dominant with nonnegative diagonal entries and has LDU factorization $\tilde{A} = \tilde{L}\tilde{D}\tilde{U}$ where $\tilde{L} = [\tilde{l}_{ij}]$, $\tilde{D} = \text{diag}(\tilde{d}_i)$, and $\tilde{U} = [\tilde{u}_{ij}]$,

$$2. |\tilde{d}_i - d_i| \leq \frac{2n\epsilon}{1-2n\epsilon} |d_i| \text{ for } i = 1, \dots, n,$$

$$3. |\tilde{u}_{ij} - u_{ij}| \leq 3n\epsilon \text{ for } i = 1, \dots, n, \text{ and}$$

$$4. \|\tilde{L} - L\|_1 \leq \frac{n(4n-1)\epsilon_0}{3(1-2n\epsilon_0)}.$$

Proof. Parts (1)-(3) are given in Theorem 2.7. For part (4), use Definition 2.3 and Theorem 2.20 with $p = i$, $q = j$ and $k = j - 1$ to show

$$\tilde{l}_{ij} = \frac{\tilde{g}_{ij}^{(j)}}{\tilde{g}_{jj}^{(j)}} = \frac{g_{ij}^{(j)} + \frac{4}{3}\chi(v_i^{(j)} + a_{ij}^{(j)})g_{j-1,j-1}^{(j-1)}}{g_{jj}^{(j)}(1 + \xi_1) \cdots (1 + \xi_j)} \quad (2.47)$$

where $|\xi_1| \leq \epsilon, \dots, |\xi_j| \leq \epsilon$, and $|\chi| \leq ((1 + \epsilon_0)^j - 1)$. Define

$$\zeta := \frac{1}{(1 + \xi_1) \cdots (1 + \xi_j)} - 1$$

and note $|\zeta| \leq \frac{1}{(1 - \epsilon)^j} - 1$. Hence,

$$\tilde{l}_{ij} = \left(l_{ij} + \frac{\frac{4}{3}\chi(v_i^{(j)} + a_{ij}^{(j)})}{a_{jj}^{(j)}} \right) (1 + \zeta) \text{ and } \tilde{l}_{ij} - l_{ij} = \zeta l_{ij} + \frac{\frac{4}{3}\chi(1 + \zeta)(v_i^{(j)} + a_{ij}^{(j)})}{a_{jj}^{(j)}}.$$

Taking the absolute value gives

$$|\tilde{l}_{ij} - l_{ij}| \leq |\zeta||l_{ij}| + \frac{4}{3}|\chi||1 + \zeta| \frac{(v_i^{(j)} + |a_{ij}^{(j)}|)}{a_{jj}^{(j)}}$$

and then summing over i yields

$$\sum_{i=j}^n |\tilde{l}_{ij} - l_{ij}| \leq |\zeta| \sum_{i=j}^n |l_{ij}| + \frac{4}{3}|\chi||1 + \zeta| \frac{\sum_{i=j}^n (v_i^{(j)} + |a_{ij}^{(j)}|)}{a_{jj}^{(j)}}.$$

By assumption that A is arranged for column diagonal dominance pivoting, which means that the produced matrix L is column diagonally dominant, we have

$$\sum_{i=j}^n |l_{ij}| \leq 1$$

for all j . Use Theorem 2.14 to obtain,

$$\begin{aligned} \sum_{i=j}^n |\tilde{l}_{ij} - l_{ij}| &\leq |\zeta| + \frac{4}{3}|\chi||1 + \zeta| \frac{\sum_{i=j}^n (v_i^{(j)} + |a_{ij}^{(j)}|)}{a_{jj}^{(j)}} \\ &\leq |\zeta| + \frac{\frac{4}{3}|\chi||1 + \zeta|(n-1)a_{jj}^{(j)}}{a_{jj}^{(j)}} \\ &= |\zeta| + \frac{4}{3}(n-1)|\chi||1 + \zeta|. \\ &\leq \left(\frac{1}{(1-\epsilon)^n} - 1 \right) + \frac{4}{3}(n-1)((1+\epsilon_0)^n - 1) \left(\frac{1}{(1-\epsilon)^n} \right) \\ &\leq \frac{n(4n-1)\epsilon_0}{3(1-2n\epsilon_0)} \end{aligned}$$

if $2n\epsilon_0 < 1$. □

Chapter 3 Inverses and Solutions to Linear Systems

Recall that two applications of the LDU factorization are inverting a matrix and solving a system of linear system. In Section 3.2 we will utilize the determinant results from the previous chapter to improve upon perturbation results for inverses and solutions to linear systems for diagonally dominant matrices. The classical perturbation bounds are normwise and dependent on the condition number. By focusing on diagonally dominant matrices, we are able to prove bounds on the inverse that are entry-wise and independent of any condition number. Similarly, for the solution to linear systems, the additional structure allows us to present bounds that are dependent on a smaller condition number. We start by presenting classical perturbation results from the literature in section 3.1.

3.1 Classical perturbation results

Consider an invertible matrix $A \in \mathbf{R}^{n \times n}$. Let $A + E \in \mathbf{R}^{n \times n}$ be a perturbation of A . Classical perturbations for the inverse are normwise and dependent on the condition number. From [9], if $\|A^{-1}\|\|E\| < 1$, then

$$\frac{\|(A + E)^{-1} - A^{-1}\|}{\|A^{-1}\|} \leq \frac{\kappa(A) \frac{\|E\|}{\|A\|}}{1 - \kappa(A) \frac{\|E\|}{\|A\|}}$$

where $\kappa(A) = \|A\|\|A^{-1}\|$ is the condition number of matrix A .

That is, the relative perturbation in the inverse of matrix A depends on the condition number of A and the relative perturbation in A . Now, suppose we have the system $Ax = b$ and consider the perturbed system $\tilde{A}\tilde{x} = \tilde{b}$. Let $\tilde{A} = A + \delta A$, $\tilde{x} = x + \delta x$, and $\tilde{b} = b + \delta b$. Then the error in the perturbed solution satisfies

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right) \quad (3.1)$$

This normwise bound expresses the relative error in the solution as a multiple (dependent on the condition number) of the relative errors in the inputs.

3.2 Relative perturbation results for diagonally dominant matrices

Recall that the entries in the inverse of matrix A can be found by computing minors of A , thus the determinant will play a crucial role in our bounds for the inverse. The proof of the following theorem, which presents a perturbation bounds for the inverse of a diagonally dominant matrix, utilizes Lemma 2.15 and Lemma 2.16 under a perturbation of type (2.7), see [13, Lemma 3 and Lemma 4].

Theorem 3.1. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$. Suppose A is nonsingular. Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ be such that*

$$|\tilde{v} - v| \leq \epsilon v \quad \text{and} \quad |\tilde{A}_D - A_D| \leq \epsilon |A_D|, \quad \text{for } j \neq i \text{ and } 0 \leq \epsilon < 1$$

Then \tilde{A} is nonsingular and

$$|(\tilde{A}^{-1})_{ij} - (A^{-1})_{ij}| \leq \frac{(3n-2)\epsilon}{1-2n\epsilon} |(A^{-1})_{jj}|.$$

Proof. From Lemma 2.15, we have

$$\det \tilde{A} = \det A(1 + \eta_1) \cdots (1 + \eta_n) \neq 0$$

with $|\eta_k| \leq \epsilon < 1$ for all k . Thus, \tilde{A} is nonsingular.

Consider $j \neq i$. Without loss of generality suppose $i = n-1$ and $j = n$. Using Lemma 2.17 with perturbation structure of type (2.7), see [13, Lemma 7], and (2.11) we have

$$\begin{aligned} |\det \tilde{A}(n', n-1') - \det A(n', n-1')| &= \left| \tilde{g}_{n-1,n}^{(n-1)} - g_{n-1,n}^{(n-1)} \right| \\ &\leq 2 \left((1 + \epsilon)^{n-1} - 1 \right) g_{n-1,n-1}^{(n-1)} \\ &\leq 2 \left((1 + \epsilon)^{n-1} - 1 \right) \det A(n', n') \end{aligned}$$

That is,

$$|\det \tilde{A}(j', i') - \det A(j', i')| \leq 2 \left((1 + \epsilon)^{n-1} - 1 \right) \det A(j', j') \quad (3.2)$$

Using Lemma 2.15, we have

$$\begin{aligned} (\tilde{A}^{-1})_{ij} - (A^{-1})_{ij} &= \frac{(-1)^{i+j} \det \tilde{A}(j', i')}{\det \tilde{A}} - \frac{(-1)^{i+j} \det A(j', i')}{\det A} \\ &= \frac{(-1)^{i+j} \det \tilde{A}(j', i')}{(\det A)(1 + \eta_1) \dots (1 + \eta_n)} - \frac{(-1)^{i+j} \det A(j', i')}{\det A} \end{aligned}$$

where $|\eta_k| \leq \epsilon$. Let $\chi = \frac{1}{(1 + \eta_1) \dots (1 + \eta_n)}$ and observe $|\chi| \leq \frac{1}{(1 - \epsilon)^n}$. Then,

$$\begin{aligned} (\tilde{A}^{-1})_{ij} - (A^{-1})_{ij} &= \frac{(-1)^{i+j} \chi (\det \tilde{A}(j', i') - \det A(j', i') + \det A(j', i'))}{(\det A)} \\ &\quad - \frac{(-1)^{i+j} \det A(j', i')}{\det A} \\ &= \frac{(-1)^{i+j} \chi (\det \tilde{A}(j', i') - \det A(j', i'))}{(\det A)} \\ &\quad + (\chi - 1) \frac{(-1)^{i+j} \det A(j', i')}{(\det A)} \end{aligned}$$

Taking the absolute value yields,

$$|(\tilde{A}^{-1})_{ij} - (A^{-1})_{ij}| = \frac{|\chi| |\det \tilde{A}(j', i') - \det A(j', i')|}{|\det A|} + |\chi - 1| \frac{|\det A(j', i')|}{|\det A|}$$

Now applying 3.2 gives,

$$\begin{aligned} |(\tilde{A}^{-1})_{ij} - (A^{-1})_{ij}| &\leq \frac{2 \left((1 + \epsilon)^{n-1} - 1 \right) |\chi| |\det A(j', j')|}{|\det A|} + |\chi - 1| \frac{|\det A(j', i')|}{|\det A|} \\ &= 2 \left((1 + \epsilon)^{n-1} - 1 \right) |\chi| |(A^{-1})_{jj}| + |\chi - 1| |(A^{-1})_{ij}| \end{aligned}$$

and hence,

$$|(\tilde{A}^{-1})_{ij} - (A^{-1})_{ij}| \leq \frac{2 \left((1 + \epsilon)^{n-1} - 1 \right)}{(1 - \epsilon)^n} |(A^{-1})_{jj}| + \left[\frac{1}{(1 - \epsilon)^n} - 1 \right] |(A^{-1})_{ij}| \quad (3.3)$$

Note that

$$\frac{1}{(1-\epsilon)^n} - 1 \leq \frac{n\epsilon}{1-n\epsilon} \quad (3.4)$$

and

$$\frac{2((1+\epsilon)^{n-1} - 1)}{(1-\epsilon)^n} \leq \frac{\frac{2(n-1)\epsilon}{1-(n-1)\epsilon}}{1 - \frac{n\epsilon}{1-n\epsilon}} \leq \frac{\frac{2(n-1)\epsilon}{1-n\epsilon}}{1 - \frac{n\epsilon}{1-n\epsilon}} = \frac{2(n-1)\epsilon}{1-2n\epsilon}.$$

Thus,

$$|(\tilde{A}^{-1})_{ij} - (A^{-1})_{ij}| \leq \left(\frac{2(n-1)\epsilon}{1-2n\epsilon} \right) |(A^{-1})_{jj}| + \left(\frac{n\epsilon}{1-n\epsilon} \right) |(A^{-1})_{ij}| \quad (3.5)$$

From [13, Theorem 1e], we have $|(A^{-1})_{ij}| \leq |(A^{-1})_{jj}|$. Substituting into (3.5) yields the desired result.

Now consider $i = j$. Using Lemma 2.16 and Lemma 2.15, we have

$$\begin{aligned} (\tilde{A}^{-1})_{ii} &= \frac{\det \tilde{A}(i', i')}{\det \tilde{A}} = \frac{(\det A(i', i'))(1 + \eta_1^{(i)}) \cdots (1 + \eta_n^{(i)})}{(\det A)(1 + \eta_1) \cdots (1 + \eta_n)} \\ &= (A^{-1})_{ii} \frac{(1 + \eta_1^{(i)}) \cdots (1 + \eta_n^{(i)})}{(1 + \eta_1) \cdots (1 + \eta_n)} = (A^{-1})_{ii}(1 + \xi)(1 + \chi) \end{aligned}$$

where $\xi = (1 + \eta_1^{(i)}) \cdots (1 + \eta_n^{(i)}) - 1$ and $\chi = \frac{1}{(1+\eta_1) \cdots (1+\eta_n)} - 1$. Observe

$$|\xi| \leq (1 + \epsilon)^n - 1 \text{ and } |\chi| \leq \frac{1}{(1 - \epsilon)^n} - 1$$

Thus, using (3.4) and (3.2) we have

$$\begin{aligned} |(\tilde{A}^{-1})_{ii} - A_{ii}^{-1}| &\leq |(A^{-1})_{ii}| |\chi| + |(A^{-1})_{ii}| |\xi| |1 + \chi| \\ &\leq |(A^{-1})_{ii}| \left(\frac{1}{(1 - \epsilon)^n} - 1 \right) + |(A^{-1})_{ii}| ((1 + \epsilon)^n - 1) \left(\frac{1}{(1 - \epsilon)^n} \right) \\ &\leq |(A^{-1})_{ii}| \frac{n\epsilon}{1 - n\epsilon} + |(A^{-1})_{ii}| \frac{n\epsilon}{1 - n\epsilon} = |(A^{-1})_{ii}| \frac{2n\epsilon}{1 - n\epsilon} \\ &\leq \frac{(3n - 2)\epsilon}{1 - 2n\epsilon} |(A^{-1})_{ii}| \end{aligned}$$

if $n \geq 2$ □

Theorem 3.1 gives that small perturbations in the data $\mathcal{D}(A_D, v)$ results in small relative perturbations in the diagonal entries of the inverse. Ideally, we would also

see small relative perturbations in the off diagonal entries, however the perturbations of the off diagonal entries depend on the corresponding diagonal entry of the inverse. This is not an artifact of our proof. Consider the matrix $A = I_2$, the 2×2 identity matrix and suppose

$$\tilde{A} = \begin{pmatrix} 1 + \epsilon & \epsilon \\ \epsilon & 1 + \epsilon \end{pmatrix}$$

for some ϵ . Note that the inverse of \tilde{A} is

$$\tilde{A}^{-1} = \frac{1}{(1 + \epsilon)^2 - \epsilon^2} \begin{pmatrix} 1 + \epsilon & -\epsilon \\ -\epsilon & 1 + \epsilon \end{pmatrix}$$

and observe that the off-diagonal entries, $\frac{-\epsilon}{1 + 2\epsilon}$, which can not be small relative to the corresponding off-diagonal entry of A^{-1} but is small relative to the corresponding diagonal entry.

The bound in Theorem 3.1 leads to some satisfactory relative normwise bounds for the inverse given below, which shows independence from any condition number.

Corollary 3.2. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$. Suppose A is nonsingular. Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ be such that*

$$|\tilde{v} - v| \leq \epsilon v \quad \text{and} \quad |\tilde{A}_D - A_D| \leq \epsilon |A_D|, \quad \text{for } j \neq i \text{ and } 0 \leq \epsilon < 1$$

Suppose $\|\cdot\|$ is either the 1-norm, the infinity norm, or the Frobenius norm. Then

$$\frac{\|\tilde{A}^{-1} - A^{-1}\|}{\|A^{-1}\|} \leq \frac{n(3n - 2)\epsilon}{1 - 2n\epsilon}$$

Proof. From [13, Theorem 1e], we have $|(A^{-1})_{ij}| \leq |(A^{-1})_{jj}| \leq \max_{k,l} |(A^{-1})_{kl}|$. Thus, Theorem 3.1 gives

$$|(\tilde{A}^{-1})_{ij} - (A^{-1})_{ij}| \leq \left(\frac{(3n - 2)\epsilon}{1 - 2n\epsilon} \right) \max_{kl} |(A^{-1})_{kl}| \quad (3.6)$$

for $i \neq j$. Similarly, from the proof of Theorem 3.1 we have

$$|(\tilde{A}^{-1})_{ii} - (A^{-1})_{ii}| \leq \left(\frac{(2n - 1)\epsilon}{1 - 2n\epsilon} \right) |(A^{-1})_{ii}| \leq \left(\frac{(2n - 1)\epsilon}{1 - 2n\epsilon} \right) \max_{kl} |(A^{-1})_{kl}|$$

If $n \geq 1$ then $3n - 2 \geq 2n - 1$ and hence

$$|(\tilde{A}^{-1})_{ii} - (A^{-1})_{ii}| \leq \left(\frac{(3n-2)\epsilon}{1-2n\epsilon} \right) \max_{kl} |(A^{-1})_{kl}|. \quad (3.7)$$

From (3.6) and (3.7) we have

$$\begin{aligned} \|\tilde{A}^{-1} - A^{-1}\| &\leq n \left(\frac{(3n-2)\epsilon}{1-2n\epsilon} \right) \max_{k,l} |(A^{-1})_{kl}| \\ &\leq n \left(\frac{(3n-2)\epsilon}{1-2n\epsilon} \right) \|A^{-1}\| \end{aligned}$$

□

With the results of Theorem 3.1 we can now present perturbation bounds for the solution to the linear system.

Theorem 3.3. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$. Suppose A is nonsingular. Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbf{R}^{n \times n}$ be such that*

$$|\tilde{v} - v| \leq \epsilon v \quad \text{and} \quad |\tilde{A}_D - A_D| \leq \epsilon |A_D|, \quad \text{for } j \neq i \text{ and } 0 \leq \epsilon < 1$$

Suppose $b, \delta b \in \mathbf{R}^{n \times n}$ such that $\|\delta b\|_\infty \leq \epsilon \|b\|_\infty$ and consider the solutions of the linear systems

$$Ax = b \quad \tilde{A}\tilde{x} = b + \delta b \quad (3.8)$$

Then, if $2n\epsilon < 1$,

$$\frac{\|\tilde{x} - x\|_2}{\|x\|_2} \leq \left[\frac{(3n^2 - 2n + 1)\epsilon + (3n^2 - 4n)\epsilon^2}{1 - 2n\epsilon} \right] \frac{\|A^{-1}\|_2 \|b\|_2}{\|x\|_2} \quad (3.9)$$

Proof. The solutions of the linear systems (3.8) can be rewritten in terms of the inverses of A and \tilde{A} as

$$x = A^{-1}b \quad \tilde{x} = \tilde{A}^{-1}b + \tilde{A}^{-1}\delta b$$

Subtracting them yields,

$$\tilde{x} - x = (\tilde{A}^{-1} - A^{-1})b + \tilde{A}^{-1}\delta b.$$

Taking the norm and applying Theorem 3.1 gives,

$$\begin{aligned}
\|\tilde{x} - x\|_2 &\leq \|\tilde{A}^{-1} - A^{-1}\|_2 \|b\|_2 + \|\tilde{A}^{-1}\|_2 \|\delta b\|_2 \\
&\leq \|\tilde{A}^{-1} - A^{-1}\|_2 \|b\|_2 + \left[\|\tilde{A}^{-1} - A^{-1}\|_2 + \|A^{-1}\|_2 \right] \epsilon \|b\|_2 \\
&\leq \frac{n(3n-2)\epsilon}{1-2n\epsilon} \|A^{-1}\|_2 \|b\|_2 + \left[\frac{n(3n-2)\epsilon}{1-2n\epsilon} \|A^{-1}\|_2 + \|A^{-1}\|_2 \right] \epsilon \|b\|_2 \\
&\leq \frac{n(3n-2)\epsilon}{1-2n\epsilon} \|A^{-1}\|_2 \|b\|_2 + \epsilon \left[\frac{n(3n-2)\epsilon}{1-2n\epsilon} + 1 \right] \|A^{-1}\|_2 \|b\|_2 \\
&\leq \left[\frac{n(3n-2)\epsilon}{1-2n\epsilon} + \epsilon \left(\frac{n(3n-2)\epsilon}{1-2n\epsilon} + 1 \right) \right] \|A^{-1}\|_2 \|b\|_2
\end{aligned}$$

Simplify and divide by $\|x\|_2$ to get the desired result. \square

The following example illustrates how Theorem 3.3 improves the classical bound in (3.1)

Example 3.4. *Let*

$$A = \begin{pmatrix} 30000 & -15000 & 15000 \\ -10000 & 20020 & 10000 \\ 20000 & 5000 & 25000 \end{pmatrix}$$

and consider the linear system $Ax = b$ where $b = [10000, 10000, -10000]^T$. Let

$$\tilde{A} = A + E = \begin{pmatrix} 30015 & -15015 & 15000 \\ -10000 & 20020 & 10000 \\ 20000 & 5000 & 25000 \end{pmatrix}$$

and suppose $(A + E)\tilde{x} = b$. The bound in (3.1) gives the relative perturbation bound

$$\frac{\|\tilde{x} - x\|_2}{\|x\|_2} \leq \frac{1351}{300} \approx 4.5033$$

compare this to the bound we presented in Theorem 3.3

$$\frac{\|\tilde{x} - x\|_2}{\|x\|_2} \leq \frac{629}{28400} \approx 2.2148 \times 10^{-2}$$

and the actual relative perturbation

$$\frac{\|\tilde{x} - x\|_2}{\|x\|_2} \leq \frac{1802}{4055777} \approx 4.4430 \times 10^{-4}.$$

In the above, we have used MAPLE to compute x and \tilde{x} exactly.

Theorem 3.3 shows that the sensitivity of the linear system $Ax = b$ to perturbations is mainly dependent upon the term $\kappa(A, b) = \|A^{-1}\|_2 \|b\|_2 / \|x\|_2$. Observe that the usual condition number $\kappa(A)$ can be defined as

$$\kappa(A) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|(A + E)^{-1} - A^{-1}\|}{\epsilon \|A^{-1}\|} : \|E\| \leq \epsilon \|A\| \right\}.$$

In practice, it is of interest to define a condition number of the linear system $Ax = b$

$$\begin{aligned} \kappa(A, b) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|(A + E)^{-1}(b + h) - A^{-1}b\|}{\epsilon \|A^{-1}b\|} \right. \\ \left. : (A + E)\tilde{x} = b + h, \frac{\|E\|}{\|A\|} \leq \epsilon, \frac{\|h\|}{\|b\|} \leq \epsilon \right\}. \end{aligned}$$

The condition number $\kappa(A, b)$ for the linear system measures the sensitivity of the solution x to relative perturbations in A and b . If we assume only b is perturbed and A is unchanged then we have

$$\begin{aligned} \kappa(A, b) &= \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|A^{-1}(b + h) - A^{-1}b\|}{\epsilon \|A^{-1}b\|} : A\tilde{x} = b + h, \|h\| \leq \epsilon \|b\| \right\} \\ &= \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|\tilde{x} - x\|}{\epsilon \|x\|} : A\tilde{x} = b + h, \|h\| \leq \epsilon \|b\| \right\} \\ &= \frac{\|A^{-1}\| \|b\|}{\|x\|} \end{aligned}$$

In general it is obvious that

$$1 \leq \kappa(A, b) \leq \kappa(A)$$

since

$$\kappa(A, b) \leq \kappa(A) + \frac{\|A^{-1}\| \|b\|}{\|A^{-1}b\|}$$

In fact, if $\kappa(A) \gg 1$, then $\kappa(A, b) \ll \kappa(A)$ for most vectors b . The condition number $\kappa(A, b)$ is usually a moderate number and, as discussed in [16], it is only large if b is

almost orthogonal to the singular vector corresponding to the smallest singular value of A .

Chapter 4 Symmetric eigenvalue problem and singular value problem

Eigenvalues play an important role in problems where the matrix is a transformation from one vector space onto itself such as in systems of linear ordinary differential equations. Singular values play an important role in problems where the matrix is a transformation from one vector space to a different vector space such as in systems of over- or underdetermined algebraic equations.

In this chapter we will discuss the classical additive perturbation bounds and the more recent work with multiplicative perturbation bounds for eigenvalues as discussed in [26]. Then we will narrow our focus to diagonally dominant matrices and discuss the work of Ye [44] that presents relative perturbation bounds for symmetric positive semidefinite diagonally dominant matrices. We then generalize this perturbation bound to all symmetric matrices. Recall that for symmetric positive definite matrices the eigenvalues and singular values coincide. Hence, we will discuss singular values of a general diagonally dominant matrix in this chapter as well.

4.1 Classical perturbation results

The classical perturbation result for the eigenvalues of the perturbed matrix $A + E$ is Weyl's Theorem given below.

Theorem 4.1 (Weyl's Theorem [9]). *Let A and E be symmetric $n \times n$ matrices. Let $\lambda_1 \leq \dots \leq \lambda_n$ be the eigenvalues of A and $\tilde{\lambda}_1 \leq \dots \leq \tilde{\lambda}_n$ be the eigenvalues of $\tilde{A} = A + E$. Then,*

$$|\lambda_i - \tilde{\lambda}_i| \leq \|E\|_2.$$

This result shows that the absolute error between the eigenvalue of the perturbed matrix $A + E$ and the corresponding eigenvalue of A is bounded by the two norm of

the absolute perturbation E . A similar Weyl-type bound is given for singular values below.

Theorem 4.2 ([9]). *Let A and E be any matrices of the same size and let $\sigma_1 \geq \dots \geq \sigma_n$ be the singular values of A and $\tilde{\sigma}_1 \geq \dots \geq \tilde{\sigma}_n$ be the singular values of $\tilde{A} = A + E$. Then,*

$$|\tilde{\sigma}_i - \sigma_i| \leq \|E\|.$$

Both additive perturbation bounds would directly produce relative perturbation bounds that are dependent on the eigenvalue or singular value itself. As we discussed in the introduction in Example 1.2 this produces several disadvantages, specifically that the bound produced for small eigenvalues and singular values could be very pessimistic. We will now consider one structure perturbation, i.e. multiplicative perturbations, which produces relative perturbation bounds.

In the rest of this section, consider a symmetric matrix A and its perturbation \tilde{A} that can be written as $\tilde{A} = D_1 A D_2$ where D_1 and D_2 are nonsingular matrices close to I . However, for the following bounds, we must assume that the perturbation \tilde{A} has the form $\tilde{A} = D A D^T$ for some nonsingular matrix D so that \tilde{A} remains symmetric.

In the following bounds, we will assume $A \in \mathbf{R}^{n \times n}$ has eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

and we will assume the perturbed matrix $\tilde{A} \in \mathbf{R}^{n \times n}$ has eigenvalues

$$\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \dots \geq \tilde{\lambda}_n$$

Theorem 4.3 (Ostrowski's Theorem [24]).

$$\lambda_{\min}[D D^T] \lambda_i \leq \tilde{\lambda}_i \leq \lambda_i \lambda_{\max}[D D^T]$$

Ostrowski's theorem bounds the ratio of exact to perturbed eigenvalues in terms of the smallest and the largest eigenvalues of DD^T , which implies the relative perturbation bound

$$\lambda_{\min}[DD^T] - 1 \leq \frac{\tilde{\lambda}_i - \lambda_i}{\lambda_i} \leq \lambda_{\max}[DD^T] - 1$$

We can rephrase Ostrowski's Theorem in terms of the two norm of DD^T .

Theorem 4.4.

$$\frac{|\lambda_i|}{\|(DD^T)^{-1}\|_2} \leq |\tilde{\lambda}_i| \leq |\lambda_i| \|DD^T\|_2$$

The term DD^T measures how close D is to being orthogonal. If $DD^T = I$ then $\lambda_{\min}[DD^*] = \lambda_{\max}[DD^T] = 1$ and $\|DD^T\|_2 = 1$. Thus, if D is close to orthogonal then the ratio $|\tilde{\lambda}_i|/|\lambda_i|$ is close to 1. We use a triangular matrix to illustrate the bound in the following example.

Example 4.5. *Consider the symmetric tridiagonal matrix*

$$A = \begin{pmatrix} 0 & \alpha_1 & & & \\ \alpha_1 & 0 & \alpha_2 & & \\ & \alpha_2 & 0 & \alpha_3 & \\ & & \alpha_3 & 0 & \alpha_4 \\ & & & \alpha_4 & 0 & \alpha_5 \\ & & & & \alpha_5 & 0 \end{pmatrix}$$

and the component-wise relative perturbation of a single off diagonal pair

$$\tilde{A} = \begin{pmatrix} 0 & \beta_1 \alpha_1 & & & \\ \beta_1 \alpha_1 & 0 & \beta_2 \alpha_2 & & \\ & \beta_2 \alpha_2 & 0 & \beta_3 \alpha_3 & \\ & & \beta_3 \alpha_3 & 0 & \beta_4 \alpha_4 \\ & & & \beta_4 \alpha_4 & 0 & \beta_5 \alpha_5 \\ & & & & \beta_5 \alpha_5 & 0 \end{pmatrix}$$

where $\beta_i \neq 0$ is close to 1. The perturbed matrix \tilde{A} can be represented as a multiplicative perturbation $\tilde{A} = DAD$, where

$$D = \begin{pmatrix} \beta_1 & & & & \\ & 1 & & & \\ & & \beta_2 & & \\ & & & \frac{\beta_3}{\beta_2} & \\ & & & & \frac{\beta_4\beta_2}{\beta_3} \\ & & & & & \frac{\beta_5\beta_3}{\beta_4\beta_2} \end{pmatrix}$$

The bound in Theorem 4.3 yields

$$\frac{1}{\eta}|\lambda_i| \leq |\tilde{\lambda}_i| \leq \eta|\lambda_i|$$

where $\eta = \prod_{j=1}^5 \max\{|\beta_j|, 1/|\beta_j|\}$. Thus, the ratio between perturbed and exact eigenvalues is close to 1 if the perturbations $|\beta_j|$ are close to one.

Following from Ostrowski's theorem, we now examine a Weyl-type bound that bounds the largest distance between a perturbed eigenvalue and the corresponding exact eigenvalue. That is, the i th largest perturbed eigenvalue is paired with the i th largest exact eigenvalue.

Theorem 4.6 ([9]).

$$|\tilde{\lambda}_i - \lambda_i| \leq |\lambda_i| \|DD^T - I\|_2.$$

If $\lambda_i \neq 0$, then we can write

$$\frac{|\tilde{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq \|DD^T - I\|_2$$

We can use Ostrowski's Theorem for eigenvalues to present similar perturbation results for singular values. In the following bounds, we will let $B \in \mathbf{R}^{m \times n}$ with singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

and we will assume the perturbed matrix $\tilde{B} = D_1 B D_2$ has singular values

$$\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \cdots \geq \tilde{\sigma}_n \geq 0$$

We obtain the following result by converting the singular value problem to an eigenvalue problem and applying the eigenvalue result in Theorem 4.3.

Theorem 4.7.

$$\frac{\sigma_i}{\|D_1^{-1}\| \|D_2^{-1}\|} \leq \hat{\sigma}_i \leq \sigma_i \|D_1\| \|D_2\|$$

If D_1 and D_2 are almost orthogonal, then the norms in Theorem 4.7 are almost 1 and thus the ratio between perturbed and exact singular value is almost 1. We illustrate the bound in Theorem 4.7 with the following example.

Example 4.8 ([26, 17]). *Consider the bidiagonal matrix*

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & & \\ & \alpha_2 & \beta_2 & \\ & & \alpha_3 & \beta_3 \\ & & & \alpha_4 \end{pmatrix}$$

and the componentwise perturbation

$$\tilde{B} = \begin{pmatrix} \gamma_1 \alpha_1 & \gamma_2 \beta_1 & & \\ & \gamma_3 \alpha_2 & \gamma_4 \beta_2 & \\ & & \gamma_5 \alpha_3 & \gamma_6 \beta_3 \\ & & & \gamma_7 \alpha_4 \end{pmatrix}$$

where $\gamma_j \neq 0$ are close to 1. We can write $\tilde{B} = D_1 B D_2$ where

$$D_1 = \begin{pmatrix} \gamma_1 & & & \\ & \frac{\gamma_1 \gamma_3}{\gamma_2} & & \\ & & \frac{\gamma_1 \gamma_3 \gamma_5}{\gamma_4} & \\ & & & \frac{\gamma_1 \gamma_3 \gamma_5 \gamma_7}{\gamma_6} \end{pmatrix}, \quad D_2 = \begin{pmatrix} 1 & & & \\ & \frac{\gamma_2}{\gamma_1} & & \\ & & \frac{\gamma_2 \gamma_4}{\gamma_3} & \\ & & & \frac{\gamma_2 \gamma_4 \gamma_6}{\gamma_5} \end{pmatrix}$$

Applying Theorem 4.7 gives

$$\frac{1}{\eta}\sigma_i \leq \tilde{\sigma}_i \leq \eta\sigma_i$$

where $\eta = \prod_{j=1}^7 \max\{|\gamma_i|, 1/|\gamma_i|\}$. Thus, if each perturbation γ_i is close to 1, then the ratio of perturbed singular values to exact singular values is close to one.

We can also convert the singular value problem to the eigenvalue problem and apply the normwise version of Ostrowski's Theorem.

Theorem 4.9.

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma_i \max\{\|I - D_1^{-1}D_1^{-T}\|, \|I - D_2^{-T}D_2^{-1}\|\}$$

The terms $\|I - D_1^{-1}D_1^{-T}\|$ and $\|I - D_2^{-T}D_2^{-1}\|$ represent the relative deviations of D_1 and D_2 from being orthogonal. Hence, if D_1 and D_2 are close to being orthogonal, then the relative error in the perturbed singular value is small.

4.2 Relative perturbation results for diagonally dominant matrices

The bounds presented in section 4.1 are only convenient for certain matrix structures. The examples we used to illustrate the bounds were tridiagonal or bidiagonal matrices. It is very challenging to write perturbations of matrices with other structures multiplicatively. In the remainder of this chapter, we will focus on symmetric diagonally dominant matrices to improve upon these results. In [44], Ye presents a relative perturbation bound for the eigenvalues of symmetric positive semidefinite diagonally dominant matrices.

Theorem 4.10 ([44]). *Let $A = [a_{ij}]$ and $\tilde{A} = [\tilde{a}_{ij}]$ be two symmetric positive semidefinite diagonally dominant matrices, and let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ and $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots \leq \tilde{\lambda}_n$ be their eigenvalues, respectively. If, for some $0 \leq \epsilon < 1$,*

$$|a_{ij} - \tilde{a}_{ij}| \leq \epsilon |a_{ij}| \text{ for all } i \neq j$$

and

$$|v_i - \tilde{v}_i| \leq \epsilon v_i \text{ for all } i,$$

where $v_i = a_{ii} - \sum_{j \neq i} |a_{ij}|$ and $\tilde{v}_i = \tilde{a}_{ii} - \sum_{j \neq i} |\tilde{a}_{ij}|$ then we have for all i

$$|\tilde{\lambda}_i - \lambda_i| \leq \epsilon \lambda_i$$

Theorem 4.10 proves that if the diagonally dominant parts, i.e. v_i , and off diagonal entries of a symmetric positive semidefinite matrix are perturbed with relative error bounded by ϵ , then the relative error in the eigenvalues is bounded by ϵ as well. This work is significant because the relative perturbation bound is independent of any condition number and the eigenvalue itself.

Positive semidefinite diagonally dominant matrices are characterized as having only non-negative diagonals. We can generalize this result by considering diagonally dominant matrix with negative diagonals. Consider the parameterization of a matrix A given in Definition 1.6. We generalize this parameterization to allow for negative diagonals below.

Definition 4.11. Let $A = [a_{ij}] \in \mathbf{R}^{n \times n}$ be any matrix. Define the following terms:

$$\begin{aligned} v_i &= |a_{ii}| - \sum_{j \neq i} |a_{ij}|, \text{ for } i = 1, \dots, n \\ A_D &= \begin{cases} 0 & \text{for } i = j \\ a_{ij} & \text{for } i \neq j \end{cases} \\ S &= \text{diag}(\text{sign}(a_{11}), \dots, \text{sign}(a_{nn})) \end{aligned}$$

Then, $A = \mathcal{D}(A_D, v, S)$ is a parameterization of A by diagonally dominant parts.

With this updated parameterization, we note that row diagonal dominance is equivalent to $v \geq 0$.

We now present perturbation results for the eigenvalues of a symmetric indefinite diagonally dominant matrix. In our proof we will construct a matrix that is nonsym-

metric with non-negative entries and apply Theorem 2.7. We also utilize the following auxiliary result.

Lemma 4.12. *Let $y \geq 0$ and $0 \leq \delta < 1$. Then,*

$$\left(\frac{1+\delta}{1-\delta}\right)^y - 1 \geq 1 - \left(\frac{1-\delta}{1+\delta}\right)^y$$

Proof. Let $x = \left(\frac{1+\delta}{1-\delta}\right)^y$ and observe $x > 0$. Thus,

$$\begin{aligned} x + \frac{1}{x} &\geq 2 \\ x - 1 &\geq 1 - \frac{1}{x} \end{aligned}$$

Substitute to obtain the desired result. □

Theorem 4.13. *Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigenvalues of the symmetric matrix $A = \mathcal{D}(A_D, v, S) \in \mathbf{R}^{n \times n}$ with $v \geq 0$. Let $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}, S)$ be a symmetric matrix such that*

$$|\tilde{v} - v| \leq \epsilon |v| \quad \text{and} \quad |\tilde{A}_D - A_D| \leq \epsilon |A_D|$$

for $0 \leq \epsilon < 1$ and let $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \dots \geq \tilde{\lambda}_n$ be the eigenvalues of \tilde{A} . If $n^3\epsilon < 1/5$, then

$$|\tilde{\lambda}_i - \lambda_i| \leq (2\nu + \nu^2)|\lambda_i|$$

for $i = 1, \dots, n$ where $\nu = \frac{4n^3\epsilon}{1 - n\epsilon}$.

Proof. Observe that S remains unperturbed. Assume A is nonsingular and has LDU decomposition $A = LDU$. Since A is symmetric then $U = L^T$, that is, $A = LDL^T$. Define the following terms

$$\begin{aligned} C &= SA, & C_D &= SA_D, & v_C &= v \\ \tilde{C} &= S\tilde{A}, & \tilde{C}_D &= S\tilde{A}_D, & \tilde{v}_C &= \tilde{v} \end{aligned}$$

Thus,

$$|\tilde{C}_D - C| \leq \epsilon |C_D| \quad \text{and} \quad |\tilde{v}_C - v_C| \leq \epsilon |v_C|$$

Note,

$$C = SA = SLDL^T = SL(SS)DL^T = (SLS)(SD)L^T$$

Thus, C has LDU factorization $C = L_C D_C U_C = (SLS)(SD)L^T$. Also note that C is a nonsymmetric diagonally dominant matrix with non-negative diagonal entries. Thus, we can apply Theorem 2.7 (1) and obtain that \tilde{C} is also a nonsingular diagonally dominant matrix with LDU factorization $\tilde{C} = \tilde{L}_C \tilde{D}_C \tilde{U}_C$. That is,

$$\tilde{C} = \tilde{L}_C \tilde{D}_C \tilde{U}_C = (S\tilde{L}S)(S\tilde{D})(\tilde{L}^T) = S\tilde{L}\tilde{D}\tilde{L}^T$$

This implies that \tilde{A} has LDU factorization $\tilde{A} = \tilde{L}\tilde{D}\tilde{L}^T$. Since A is square and invertible by assumption, then this LDU factorization is unique. Furthermore, from parts (2) and (3) of Theorem 2.7 we obtain the following bounds.

$$(\tilde{D}_C)_{ii} = (D_C)_{ii} \frac{(1 + \eta_{i1}) \cdots (1 + \eta_{ii})}{(1 + \eta_{i-1,1}) \cdots (1 + \eta_{i-1,i-1})}, \text{ where } |\eta_{ik}| \leq \epsilon, |\eta_{i-1,p}| \leq \epsilon \quad (4.1)$$

for $k = 1, \dots, i$ and $p = 1, \dots, i-1$, and

$$|(\tilde{L}_C)_{ij} - (L_C)_{ij}| \leq 3i\epsilon \leq 3n\epsilon \quad (4.2)$$

From (4.1), we obtain

$$\tilde{d}_{ii} = d_{ii} \frac{(1 + \eta_{i1}) \cdots (1 + \eta_{ii})}{(1 + \eta_{i-1,1}) \cdots (1 + \eta_{i-1,i-1})}$$

Set $\gamma_i = \sqrt{\frac{(1 + \eta_{i1}) \cdots (1 + \eta_{ii})}{(1 + \eta_{i-1,1}) \cdots (1 + \eta_{i-1,i-1})}} - 1$. Then, $\tilde{d}_{ii} = d_{ii}(1 + \gamma_i)^2$ and we can use Lemma 4.12 to obtain

$$|\gamma_i| \leq \left(\frac{1 + \epsilon}{1 - \epsilon} \right)^{n/2} - 1.$$

It can be shown by induction that

$$\left(\frac{1 + \epsilon}{1 - \epsilon} \right)^{n/2} - 1 \leq \frac{n\epsilon}{1 - n\epsilon}$$

Set $W = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_n)$. Then, we can write \tilde{D} as

$$\tilde{D} = (I + W)D(I + W) \text{ with } \|W\|_{\max} \leq \frac{n\epsilon}{1 - n\epsilon}. \quad (4.3)$$

Also, note

$$\tilde{L}_C - L_C = S\tilde{L}S - SLS = S(\tilde{L} - L)S = \tilde{L} - L$$

and thus, using (4.2) we have

$$|\tilde{l}_{ij} - l_{ij}| \leq 3n\epsilon \text{ and } \|\Delta_L\|_{\max} \leq 3n\epsilon. \quad (4.4)$$

Consider

$$\begin{aligned} \tilde{A} &= \tilde{L}\tilde{D}\tilde{L}^T = (L + \Delta_L)(I + W)D(I + W)(L + \Delta_L)^T \\ &= [L + \Delta_L + LW + \Delta_L W] D [L + \Delta_L + LW + \Delta_L W]^T \\ &= (I + F)LWL^T(I + F)^T \end{aligned}$$

where $F = \Delta_L L^{-1} + LWL^{-1} + \Delta_L WL^{-1}$. Since L is column diagonally dominant, then $\|L\|_{\max} \leq 1$. Moreover, the inverse of a column diagonally dominant matrix is row diagonally dominant, so $\|L^{-1}\|_{\max} \leq 1$. These bounds, combined with (4.3) and (4.4) yield

$$\begin{aligned} \|F\|_2 &\leq n\|F\|_{\max} = n\|\Delta_L L^{-1} + LWL^{-1} + \Delta_L WL^{-1}\|_{\max} \\ &\leq n[\|\Delta_L\|_{\max}\|L^{-1}\|_{\max} + \|LW\|_{\max}\|L^{-1}\|_{\max} + \|\Delta_L\|_{\max}\|WL^{-1}\|_{\max}] \\ &\leq n\left[3n\epsilon + \frac{n\epsilon}{1 - n\epsilon} + 3n\epsilon\left(\frac{n\epsilon}{1 - n\epsilon}\right)\right] = \frac{4n^2\epsilon}{1 - n\epsilon} \end{aligned}$$

If we restrict $n^2\delta < 1/5$, then $\|F\| < 1$ which implies $I + F$ is nonsingular. Hence, we can apply [17, Theorem 2.1], which states if $\tilde{A} = D^T A D$ for some nonsingular matrix D , then

$$|\tilde{\lambda}_i - \lambda_i| \leq |\lambda_i| \|D^T D - I\|, \quad (4.5)$$

with $D = (I + F)^T$. Note that

$$\begin{aligned}
\|D^T D - I\| &= \|(I + F)(I + F)^T - I\| = \|F + F^T + FF^T\| \\
&\leq \|F\| + \|F\| + \|F\|\|F\| = 2\|F\| + \|F\|^2 \\
&\leq 2\left(\frac{4n^2\epsilon}{1 - n\epsilon}\right) + \left(\frac{4n^2\epsilon}{1 - n\epsilon}\right)^2 = 2\nu + \nu^2
\end{aligned}$$

where $\nu = \left(\frac{4n^2\epsilon}{1 - n\epsilon}\right)$. Hence, from (4.5) becomes

$$|\tilde{\lambda}_i - \lambda_i| \leq |\lambda_i| (2\nu + \nu^2),$$

□

Note that this result holds if the matrix A is singular. To prove, from [13, Theorem 2] there exists a permutation matrix P such that $B = PAP^T$ has an LDU factorization and we can construct a parameterization of B from the parameterization of $A = \mathcal{D}(A_D, v, S)$. From here, we can proceed with the proof using B .

We now consider the singular values for nonsymmetric diagonally dominant matrices with nonnegative diagonals. Recall that for symmetric matrices, the singular values are the absolute values of the eigenvalues and thus we can use bounds previously. In the following theorem we show that if a diagonally dominant matrix is perturbed such that the diagonally dominant parts and off diagonal entries have relative errors bounded by some ϵ , then the singular values have relative errors that depend on only ϵ and the size of the matrix.

Theorem 4.14. *Let $A = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that $v \geq 0$ and let $\tilde{A} = \mathcal{D}(A_D, v) \in \mathbf{R}^{n \times n}$ be such that*

$$|\tilde{v} - v| \leq \epsilon v \text{ and } |\tilde{A}_D - A_D| \leq \epsilon |A_D|$$

for some $0 \leq \epsilon < \frac{1}{3}$. Let $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ and $\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \dots \geq \tilde{\sigma}_n$ be the singular values of A and \tilde{A} , respectively. Define

$$\nu := \frac{n^3(4n - 1)\epsilon_0}{1 - 2n\epsilon_0}$$

where $\epsilon_0 = \frac{12\epsilon}{1-\epsilon}$. If $2n\epsilon_0 < 1$ and $\nu < 1$, then

$$|\tilde{\sigma}_i - \sigma_i| \leq (2\nu + \nu^2)\sigma_i$$

for $i = 1, \dots, n$.

Proof. Without loss of generality, we assume that A is arranged for column diagonal dominance pivoting. Note that A has LDU factorization $A = LDU$ and apply Theorem 2.7 to show that \tilde{A} has LDU factorization $\tilde{A} = \tilde{L}\tilde{D}\tilde{U}$ such that for $2n\epsilon < 1$,

$$\tilde{d}_i = d_i(1 + w_i) \text{ where } |w_i| \leq \frac{2n\epsilon}{1 - 2n\epsilon} \quad (4.6)$$

which gives

$$\tilde{D} = D(I + W) \text{ with } W = \text{diag}(w_1, w_2, \dots, w_n),$$

and,

$$\|\Delta_U\|_{\max} = 3n\epsilon \text{ where } \Delta_U := \tilde{U} - U \quad (4.7)$$

From Theorem 2.21 we have

$$\|\Delta_L\|_2 = \frac{n^{3/2}(4n-1)\epsilon_0}{3(1-2n\epsilon_0)} \text{ where } \Delta_L := \tilde{L} - L \quad (4.8)$$

Observe

$$\begin{aligned} \tilde{A} &= \tilde{L}\tilde{D}\tilde{U} = (L + \Delta_L)D(I + W)(U + \Delta_U) \\ &= (I + \Delta_L L^{-1})LD(U + \Delta_U + WU + W\Delta_U) \\ &= (I + \Delta_L L^{-1})LDU(I + U^{-1}\Delta_U + U^{-1}WU + U^{-1}W\Delta_U) \\ &= (I + E)LDU(I + F) \end{aligned}$$

where

$$E := \Delta_L L^{-1}, \text{ and} \quad (4.9)$$

$$F := U^{-1}\Delta_U + U^{-1}WU + U^{-1}W\Delta_U. \quad (4.10)$$

Since L is column diagonally dominant, then L^{-1} is row diagonally dominant and hence,

$$\|L\|_{\max} \leq 1 \text{ and } \|L^{-1}\|_{\max} \leq 1. \quad (4.11)$$

Similarly, since U is row diagonally dominant, then U^{-1} is column diagonally dominant. Hence,

$$\|U\|_{\max} \leq 1 \text{ and } \|U^{-1}\|_{\max} \leq 1. \quad (4.12)$$

From (4.6), we have

$$\|W\|_{\max} \leq \frac{2n\epsilon}{1 - 2n\epsilon}$$

and thus,

$$\|WU\|_{\max} \leq \|W\|_{\max}\|U\|_{\max} \leq \frac{2n\epsilon}{1 - 2n\epsilon}, \text{ and} \quad (4.13)$$

$$\|U^{-1}W\|_{\max} \leq \|U^{-1}\|_{\max}\|W\|_{\max} \leq \frac{2n\epsilon}{1 - 2n\epsilon} \quad (4.14)$$

Combining (4.7), (4.13), (4.14), and (4.12) with (4.10) yields

$$\begin{aligned} \|F\|_{\max} &\leq \|U^{-1}\|_{\max}\|\Delta_U\|_{\max} + \|U^{-1}\|_{\max}\|WU\|_{\max} + \|U^{-1}W\|_{\max}\|\Delta_U\|_{\max} \\ &\leq 3n\epsilon + \frac{2n\epsilon}{1 - 2n\epsilon} + \frac{2n\epsilon}{1 - 2n\epsilon}(3n\epsilon) \\ &= \frac{5n\epsilon}{1 - 2n\epsilon} \end{aligned}$$

Thus, we have

$$\|F\|_2 \leq n\|F\|_{\max} \leq \frac{5n^2\epsilon}{1 - 2n\epsilon} \leq \nu$$

From (4.9) and (4.8)

$$\|E\|_2 \leq \|\Delta_L\|_2\|L^{-1}\|_2 \leq \left(\frac{n^{3/2}(4n-1)\epsilon_0}{3(1-2n\epsilon_0)} \right) (n) = \nu$$

Thus, if $\nu \leq 1$ then both $(I + E)$ and $(I + F)$ are nonsingular. Therefore we can apply [17, Theorem 3.3] to obtain

$$|\tilde{\sigma}_i - \sigma_i| \leq \gamma\sigma_i \quad (4.15)$$

for $i = 1, 2, \dots, n$ where $\gamma = \max\{\|(I + E)(I + E)^T - I\|_2, \|(I + F)^T(I + F) - I\|_2\}$.

Note that

$$\begin{aligned} \|(I + E)(I + E)^T - I\|_2 &= \|I + E + E^T + EE^T - I\|_2 = \|E + E^T + EE^T\|_2 \\ &\leq \|E\|_2 + \|E^T\|_2 + \|E\|\|E^T\|_2 \leq 2\|E\|_2 + \|E\|_2^2 \\ &\leq 2\nu + \nu^2 \end{aligned}$$

Similarly,

$$\|(I + F)(I + F)^T - I\|_2 \leq 2\nu + \nu^2$$

Thus, $\gamma = 2\nu + \nu^2$. Combining with (4.15) yields

$$|\tilde{\sigma}_i - \sigma_i| \leq (2\nu + \nu^2)\sigma_i \tag{4.16}$$

□

Chapter 5 Nonsymmetric Eigenvalue Problem

In the previous chapter we discussed perturbation bounds for symmetric matrices. In this chapter, we will consider the nonsymmetric case. The nonsymmetric eigenvalue problem is generally much more complex than the symmetric eigenvalue problem. Recall that λ is called an eigenvalue of A with corresponding (right) eigenvector x if $Ax = \lambda x$. We can also define a left eigenvector of A . That is, λ is an eigenvalue of A with corresponding left eigenvector y if $y^*A = \lambda y^*$. In the symmetric case, a left eigenvector is also a right eigenvector of A . This is not necessarily true for the nonsymmetric case. Moreover, in the nonsymmetric case the eigenvalues of A can be complex. Most algorithms that compute the eigenvalues of a nonsymmetric A use a similarity transformation to transform A into a canonical form for which it is easier to compute its eigenvalues and eigenvectors of A . Ideally, the canonical form would be a triangular matrix, since the eigenvalues of a triangular matrix are simply the diagonal entries. The simplest canonical form under similarity transformation is the Jordan canonical form.

Theorem 5.1 (Jordan canonical form [9]). *Given A , there exists a nonsingular matrix S such that $S^{-1}AS = J$, where J is in Jordan canonical form. This means that J is block diagonal, with $J = \text{diag}(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \dots, J_{n_k}(\lambda_k))$ and*

$$J_{n_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in \mathbf{C}^{n_i \times n_i}.$$

J is unique, up to permutations of its diagonal blocks.

While theoretically useful, the Jordan canonical form is hard to compute in a

numerically stable fashion. This means that any small perturbation can change it completely. Consider the following example from [9]

Example 5.2. *Let*

$$J_n(0) = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},$$

For an arbitrarily small ϵ , adding $k\epsilon$ to the (k, k) entry changes the eigenvalues to the n distinct values $k \cdot \epsilon$ and so the Jordan form changes from $J_n(0)$ to $\text{diag}(\epsilon, 2\epsilon, \dots, n\epsilon)$.

For this reason, the Jordan canonical form can not be computed stably. Indeed, the Jordan canonical form can not be computed in a backward stable manner.

Example 5.3 ([9]). *Suppose $S^{-1}AS = J$ exactly where S is ill conditioned, i.e. $\kappa(S) \gg 1$. Suppose we manage to compute S exactly and J with a tiny error ΔJ , where $\|\Delta J\| \leq \mathcal{O}(\epsilon)\|A\|$. Now suppose we try to bound the backward error. That is, we want to determine how big ΔA must be so that $S^{-1}(A + \Delta A)S = J + \Delta J$. This gives $\Delta A = S\Delta JS^{-1}$ and we can only conclude that $\|\Delta A\| \leq \|S\|\|\Delta J\|\|S^{-1}\| = \mathcal{O}(\epsilon)\kappa(S)\|A\|$. Thus, $\|\Delta A\|$ may be much larger than $\epsilon\|A\|$, which prevents backward stability.*

Example 5.4 ([9]). *Let*

$$A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \epsilon & & & 0 \end{pmatrix}.$$

The characteristic polynomial for A is $\lambda^n - \epsilon = 0$. For $\epsilon > 0$, one real eigenvalue is $\lambda = \lambda(\epsilon) = \sqrt[n]{\epsilon}$. Note that

$$\frac{d(\lambda(\epsilon) - \lambda(0))}{d\epsilon} = \frac{\epsilon^{\frac{1}{n}-1}}{n}$$

which tends to infinity as ϵ tends to 0. Thus, the eigenvalue perturbation $\lambda(\epsilon) - \lambda(0)$ may be infinitely larger than a small change ϵ .

In Section 5.1, we will present classical perturbation bounds from the literature. These bounds will require A to be diagonalizable. That is, there is a similarity transformation X such that $A = X\Lambda X^{-1}$ where Λ is a diagonal matrix. We will show that these bounds can be very pessimistic if A is close to having multiple eigenvalues with Jordan blocks. In section 5.2, we will consider only simple eigenvalues and will use rank revealing decompositions to improve upon the classical bounds.

5.1 Classical perturbation results

In section 4.1, we presented Weyl type perturbation bounds for the eigenvalues of symmetric matrices. In this section, we will focus on Bauer-Fike type bounds. Bauer-Fike type bounds are two-norm bounds on the distance between a perturbed eigenvalue and the closest exact eigenvalue.

Theorem 5.5 (Bauer-Fike Theorem [26]). *Let $A \in \mathbf{R}^{n \times n}$ have eigendecomposition $A = X\Lambda X^{-1}$ where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ and let $\tilde{A} = A + E$ be a perturbation of A . If $\tilde{\lambda}$ is an eigenvalue of \tilde{A} , then*

$$\min_i |\lambda_i - \tilde{\lambda}| \leq \kappa(X) \|E\|_2. \quad (5.1)$$

If A is nonsingular,

$$\min_i \frac{|\lambda_i - \tilde{\lambda}|}{|\lambda_i|} \leq \kappa(X) \|A^{-1}E\|_2 \quad (5.2)$$

The bound in (5.1) produces the following relative bound

$$\frac{|\lambda_i - \tilde{\lambda}|}{|\lambda_i|} \leq \kappa(X) \frac{\|E\|_2}{|\lambda_i|} \leq \kappa(X) \|A^{-1}\|_2 \|E\|_2 \leq \kappa(X) \kappa(A) \frac{\|E\|_2}{\|A\|_2}$$

which is very close to the relative error bound (5.2). Thus, the relative error in the eigenvalues depends on how close the eigenvalues are to being multiple eigenvalues

with a Jordan block as measured by $\kappa(X)$, how ill-conditioned A is as measured by $\kappa(A)$, and in the relative perturbation $\|E\|_2/\|A\|_2$.

One disadvantage to the bounds in Theorem 5.5 is that they can be very pessimistic, especially if an eigenvalue is close to being multiple with Jordan block. Consider the following example.

Example 5.6. *Let*

$$A = \begin{pmatrix} 1 & 1001 \\ 10^{-3} & 1 \end{pmatrix}$$

and the perturbed matrix \tilde{A}

$$\tilde{A} = A + E = \begin{pmatrix} 1 & 1001 \\ 0 & 1 \end{pmatrix}.$$

The eigenvalues of A are $\lambda = 1 \pm \frac{\sqrt{10010}}{100} \approx 1 \pm 1.0005$ with eigenvector matrix

$$X = \begin{pmatrix} 10\sqrt{10010} & -10\sqrt{10010} \\ 1 & 1 \end{pmatrix}.$$

The eigenvalue of \tilde{A} is $\lambda = 1$. The actual absolute and relative errors are

$$\min_i |\lambda_i - \tilde{\lambda}| = \frac{\sqrt{10010}}{100} \approx 1.0005,$$

and

$$\min_i \frac{|\lambda_i - \tilde{\lambda}|}{|\lambda_i|} = 5.001 \times 10^{-1}.$$

However, the bounds in Theorem 5.5 give

$$\min_i |\lambda_i - \tilde{\lambda}| \leq 1.0015,$$

and

$$\min_i \frac{|\lambda_i - \tilde{\lambda}|}{|\lambda_i|} \leq 1.0025 \times 10^6$$

Now, suppose the perturbation of A can be written multiplicatively. That is, suppose \tilde{A} has the form $\tilde{A} = D_1 A D_2$ where D_1, D_2 are non-singular matrices close to I . Note that if $D_2 = D_1^{-1}$ then \tilde{A} is a similarity transformation of A and thus has the same eigenvalues as A . If A is Hermitian and $D_2 = D_1^*$, then \tilde{A} is a congruency transformation of A , and thus has the same inertia as A . That is, A and \tilde{A} have the same number of positive, negative, and zero eigenvalues.

Theorem 5.7. *If A is diagonalizable and $\tilde{A} = D_1 A D_2$ where D_1 and D_2 are nonsingular, then*

$$\min_i |\lambda_i - \hat{\lambda}| \leq |\hat{\lambda}| \kappa(X) \|I - D_1^{-1} D_2^{-1}\|$$

The term $\|I - D_1^{-1} D_2^{-1}\|_2$ represents a relative deviation from similarity, that is, the relative perturbation of eigenvalues is small when \hat{A} is close to being a similarity transformation of A . This bound is tight if $D_2 = D_1^{-1}$ or if $\hat{\lambda} = 0$.

As in the Bauer-Fike type bounds, the terms $\|D_1 - D_2^{-1}\|$ and $\|D_2 - D_1^{-1}\|$ represent a relative deviation from similarity. That is, the relative error bound is small only if D_1 and D_2 are close to a similarity transformation.

The above theorems bound all the eigenvalues together. It turns out for non-symmetric matrices, different eigenvalues may have different perturbation properties. The following perturbation analysis more precisely describes the perturbation for each eigenvalue.

Theorem 5.8 ([9]). *Let λ be a simple eigenvalue of $A \in \mathbf{R}^{n \times n}$ with right eigenvector x and left eigenvector y . Let $\tilde{\lambda}$ be the eigenvalue of $\tilde{A} = A + E$ that is closest to λ , where $E \in \mathbf{R}^{n \times n}$. Then*

$$\tilde{\lambda} - \lambda = \frac{y^* E x}{y^* x} + \mathcal{O}(\|E\|_2^2) \quad (5.3)$$

or,

$$|\tilde{\lambda} - \lambda| \leq \sec \theta(y, x) \|E\|_2 + \mathcal{O}(\|E\|_2^2) \quad (5.4)$$

where $\theta(y, x)$ is the acute angle between x and y .

The term $\sec \theta(y, x) = \frac{\|y^*\|_2 \|x\|_2}{|y^* x|}$ is the condition number of the eigenvalue λ . It describes the sensitivity of λ to perturbation. The perturbation bound in Theorem 5.8 leads to the relative perturbation bound

$$\frac{|\tilde{\lambda} - \lambda|}{|\lambda|} \leq \frac{\sec \theta(y, x) \|E\|_2}{|\lambda|}. \quad (5.5)$$

5.2 Relative perturbation bounds

The relative perturbation bound in (5.5) depends on both the condition number of the eigenvalue and the eigenvalue itself. For diagonally dominant matrices, present in this section a new bound that removes the dependence on the eigenvalue.

We will first use a rank-revealing decomposition and then, more specifically, the LDU factorization to derive relative perturbation bounds. Theorem 5.8 provides an absolute error bound for a perturbed eigenvalue. Suppose we have any matrix A and we perturb it slightly to form $\tilde{A} = A + E$. Then the error between the corresponding eigenvalues of A and \tilde{A} can be bounded in terms of the perturbation in A and the condition number of the eigenvalue. This result depends on the left and right eigenvectors of A . We first present a variation of Theorem 5.8

Lemma 5.9. *Let λ be a simple eigenvalue of $A \in \mathbf{R}^{n \times n}$ with right eigenvector x . Let $\tilde{\lambda}$ be the corresponding eigenvalue of $\tilde{A} = A + E$, where $E \in \mathbf{R}^{n \times n}$, with left eigenvector \tilde{y} . Then,*

$$\tilde{\lambda} - \lambda = \frac{\tilde{y}^* E x}{\tilde{y}^* x} \quad (5.6)$$

or

$$|\tilde{\lambda} - \lambda| \leq \sec \theta(\tilde{y}^*, x) \|E\|_2. \quad (5.7)$$

Proof. Since $E = \tilde{A} - A$, then using the definition of eigenvalue we have

$$\tilde{y}^* E x = \tilde{y}^* (\tilde{A} - A) x = \tilde{y}^* \tilde{A} x - \tilde{y}^* A x = (\tilde{y}^* \tilde{\lambda}) x - \tilde{y}^* (\lambda x) = (\tilde{\lambda} - \lambda) \tilde{y}^* x.$$

from which (5.6) follows. \square

Notice the result in (5.7) is very similar to the result in (5.4), however one advantage is the lack of higher order terms. On the other hand, it depends on the left eigenvector \tilde{y} of \tilde{A} , which is less desirable in general. However, this is advantageous for our purpose as it will become evident in the proof of the following theorem, in which we show that given a matrix A , we can compute its eigenvalues with high relative accuracy provided we can find an accurate rank revealing decomposition.

Theorem 5.10. *Let $A = XDY^T \in \mathbf{R}^{n \times n}$ be a rank revealing decomposition and let $\tilde{A} = \tilde{X}\tilde{D}\tilde{Y}^T \in \mathbf{R}^{n \times n}$ where \tilde{X} , \tilde{D} , and \tilde{Y} are defined as follows:*

$$\tilde{X} = X + \Delta_X, \quad \tilde{D} = D + \Delta_D, \quad \tilde{Y} = Y + \Delta_Y \quad (5.8)$$

where

$$|(\Delta_D)_{ii}| \leq \epsilon |D_{ii}|, \quad \|\Delta_X\|_2 \leq \epsilon \|X\|_2, \quad \text{and} \quad \|\Delta_Y\|_2 \leq \epsilon \|Y\|_2$$

with $0 \leq \kappa\epsilon < 1$ where $\kappa = \max\{\kappa_2(X), \kappa_2(Y)\}$. Let λ be a simple eigenvalue of A and $\tilde{\lambda}$ be the corresponding eigenvalue of \tilde{A} . Then,

$$\frac{|\tilde{\lambda} - \lambda|}{|\lambda|} \leq \epsilon \kappa \sec \theta(\tilde{y}, x) \frac{1 + (1 + \epsilon)}{1 - \epsilon \kappa \sec \theta(\tilde{y}, x) \gamma} \quad (5.9)$$

$$\text{where } \gamma = \frac{1 + \epsilon}{(1 - \epsilon)(1 - \epsilon \kappa)} = 1 + \mathcal{O}(\epsilon)$$

Proof. Observe

$$\begin{aligned} \tilde{A} - A &= \tilde{X}\tilde{D}\tilde{Y}^T - A = \tilde{X}(D + \Delta_D)(Y^T + \Delta_{Y^T}) - A \\ &= \tilde{X}DY^T + \tilde{X}D\Delta_{Y^T} + \tilde{X}\Delta_D Y^T + \tilde{X}\Delta_D \Delta_{Y^T} - A \\ &= (X + \Delta_X)DY^T + \tilde{X}D\Delta_{Y^T} + \tilde{X}\Delta_D Y^T + \tilde{X}\Delta_D \Delta_{Y^T} - XDY^T \\ &= \Delta_X DY^T + \tilde{X}D\Delta_{Y^T} + \tilde{X}\Delta_D Y^T + \tilde{X}\Delta_D \Delta_{Y^T} \end{aligned}$$

Applying (5.6) yields

$$(\tilde{y}^* x) (\tilde{\lambda} - \lambda) = \tilde{y}^* \Delta_X DY^T x + \tilde{y}^* \tilde{X} \Delta_D U x + \tilde{y}^* \tilde{X} D \Delta_{Y^T} x + \tilde{y}^* \tilde{X} \Delta_D \Delta_{Y^T} x$$

Taking the absolute value, we have

$$\begin{aligned}
|\tilde{y}^* x| \left| \tilde{\lambda} - \lambda \right| &\leq |\lambda| \|\tilde{y}^*\|_2 \|\Delta_X\|_2 \|X^{-1}\|_2 \|x\|_2 + |\lambda| \|\tilde{y}^*\|_2 \|\tilde{X}\|_2 \|\Delta_D D^{-1}\|_2 \|X^{-1}\|_2 \|x\|_2 \\
&\quad + |\tilde{\lambda}| \|\tilde{y}^*\|_2 \|\tilde{Y}^{-T}\|_2 \|\tilde{D}^{-1} D\|_2 \|\Delta_{Y^T}\|_2 \|x\|_2 \\
&\quad + |\tilde{\lambda}| \|\tilde{y}^*\|_2 \|\tilde{Y}^{-T}\|_2 \|\tilde{D}^{-1} \Delta_D\|_2 \|\Delta_{Y^T}\|_2 \|x\|_2
\end{aligned} \tag{5.10}$$

Since $|(\Delta_D D^{-1})_{ii}| \leq |(\Delta_D)_{ii}| |D_{ii}^{-1}| \leq \epsilon$ then

$$\|\Delta_D D^{-1}\|_2 \leq \epsilon \tag{5.11}$$

from which it follows

$$\|\tilde{D}^{-1} D\|_2 = \|(I + D^{-1} \Delta_D)^{-1}\|_2 \leq \frac{1}{1 - \epsilon} \tag{5.12}$$

and

$$\|\tilde{D}^{-1} \Delta_D\|_2 = \|(I + D^{-1} \Delta_D)^{-1} D^{-1} \Delta_D\|_2 \leq \frac{\epsilon}{1 - \epsilon}. \tag{5.13}$$

By assumption, we have

$$\|\Delta_X\|_2 \|X^{-1}\|_2 \leq \epsilon \|X\|_2 \|X^{-1}\|_2 = \epsilon \kappa_2(X) \leq \epsilon \kappa \tag{5.14}$$

which yields

$$\begin{aligned}
\|\tilde{X}\|_2 \|X^{-1}\|_2 &\leq (\|X\|_2 + \|\Delta_X\|_2) \|X^{-1}\|_2 = \|X\|_2 \|X^{-1}\|_2 + \|\Delta_X\|_2 \|X^{-1}\|_2 \\
&\leq \kappa_2(X) + \epsilon \kappa_2(X) = (1 + \epsilon) \kappa_2(X) \leq (1 + \epsilon) \kappa.
\end{aligned} \tag{5.15}$$

Also, by assumption, we have

$$\|\Delta_Y\|_2 \|Y^{-T}\|_2 \leq \epsilon \|Y\|_2 \|Y^{-1}\|_2 \leq \epsilon \kappa_2(Y) \leq \epsilon \kappa \tag{5.16}$$

and thus

$$\begin{aligned}
\|\tilde{Y}^{-T}\|_2 \|\Delta_Y\|_2 &\leq \|(I + Y^{-1} \Delta_Y)^{-1} Y^{-1}\|_2 \|\Delta_Y\|_2 \\
&\leq \frac{1}{1 - \epsilon \kappa_2(Y)} \|Y^{-1}\|_2 \epsilon \|Y\|_2 = \frac{\epsilon \kappa_2(Y)}{1 - \epsilon \kappa_2(Y)} \leq \frac{\epsilon \kappa}{1 - \epsilon \kappa}
\end{aligned} \tag{5.17}$$

Substituting the bounds (5.11), (5.12), (5.13), (5.14), (5.15), and (5.17) into (5.10) gives

$$\begin{aligned}
|\tilde{y}^* x| |\tilde{\lambda} - \lambda| &\leq |\lambda| \|\tilde{y}^*\|_2 \epsilon \kappa \|x\|_2 + |\lambda| \|\tilde{y}^*\|_2 \epsilon (1 + \epsilon) \kappa \|x\|_2 \\
&\quad + |\tilde{\lambda}| \|\tilde{y}^*\|_2 \left(\frac{\epsilon \kappa}{1 - \epsilon \kappa} \right) \left(\frac{1}{1 - \epsilon} \right) \|x\|_2 \\
&\quad + |\tilde{\lambda}| \|\tilde{y}^*\|_2 \left(\frac{\epsilon \kappa}{1 - \epsilon \kappa} \right) \left(\frac{\epsilon}{1 - \epsilon} \right) \|x\|_2
\end{aligned}$$

Rearranging yields

$$|\tilde{\lambda} - \lambda| \leq |\lambda| (\epsilon \kappa (1 + (1 + \epsilon)) \sec \theta(\tilde{y}, x)) + |\tilde{\lambda}| \frac{\epsilon \kappa}{(1 - \epsilon)(1 - \epsilon \kappa)} \sec \theta(\tilde{y}, x) \epsilon.$$

Replace $|\tilde{\lambda}|$ with $|\tilde{\lambda} - \lambda + \lambda|$ and use the triangle inequality to obtain the desired result. \square

Notice that our bound presented in Theorem 5.10 still depends on the condition number of the eigenvalue λ , $\sec \theta(\tilde{y}, x)$, but it is independent of the eigenvalue itself. Traditional eigenvalue perturbation bounds are dependent upon the eigenvalues. In the following sections we will utilize the structure of diagonally dominant matrices to provide a similar perturbation bound.

Now consider the LDU factorization of a diagonally dominant matrix. In the following theorem, the bounds presented in Theorem 2.21 are used to obtain a relative perturbation bound for the eigenvalues.

Theorem 5.11. *Let λ be a simple eigenvalue of $A = \mathcal{D}(A_D, v) \in \mathbb{R}^{n \times n}$ where $v \geq 0$ with right eigenvector x . Let $\tilde{\lambda}$ be an eigenvalue of $\tilde{A} = \mathcal{D}(\tilde{A}_D, \tilde{v}) \in \mathbb{R}^{n \times n}$ with left eigenvector \tilde{y} such that $\tilde{y}^* x \neq 0$ and*

$$|\tilde{v} - v| \leq \epsilon v \text{ and } |\tilde{A}_D - A_D| \leq \epsilon |A_D|, \text{ for some } 0 \leq \epsilon < \frac{1}{3}. \quad (5.18)$$

Then, if $7n^3\epsilon_0 \leq 1$, we have

$$\frac{|\tilde{\lambda} - \lambda|}{|\lambda|} \leq \sec \theta(\tilde{y}, x) \frac{(13n^4\epsilon_0 - n^3\epsilon_0 - 12n^5\epsilon^2 - 3n^4\epsilon^2)\gamma}{1 + \epsilon \sec \theta(y, x) 3n^3\gamma} \quad (5.19)$$

where $\gamma = \frac{1}{1 - 7n^3\epsilon_0}$, $\epsilon_0 = \frac{6\epsilon}{1 - \epsilon}$, and $\sec \theta(\tilde{y}, x) = \frac{\|\tilde{y}^*\|_2 \|x\|_2}{|\tilde{y}^* x|}$.

Proof. Suppose A is arranged for column diagonal dominance pivoting and has LDU factorization $A = LDU$. From Theorem 2.21(1) we have $\tilde{A} = \tilde{L}\tilde{D}\tilde{U}$. Define Δ_L , Δ_D , and Δ_U such that

$$\tilde{L} = L + \Delta_L, \quad \tilde{D} = D + \Delta_D, \quad \tilde{U} = U + \Delta_U$$

Then, we can write $E = \tilde{A} - A$ as

$$E = \Delta_L DU + \tilde{L} \Delta_D U + \tilde{L} D \Delta_U + \tilde{L} \Delta_D \Delta_U$$

as done in the proof of Theorem 5.10. Use Lemma 5.9, to obtain

$$\begin{aligned} (\tilde{\lambda} - \lambda)(\tilde{y}^* x) &= \lambda \tilde{y}^* \Delta_L L^{-1} x + \lambda \tilde{y}^* \tilde{L} \Delta_D D^{-1} L^{-1} x \\ &\quad + \tilde{\lambda} \tilde{y}^* \tilde{U}^{-1} \tilde{D}^{-1} D \Delta_U x + \tilde{\lambda} \tilde{y}^* \tilde{U}^{-1} \tilde{D}^{-1} \Delta_D \Delta_U x. \end{aligned} \quad (5.20)$$

From Theorem 2.21 we have

$$\|\Delta_D D^{-1}\|_2 \leq \frac{2n\epsilon}{1 - 2n\epsilon} \quad (5.21)$$

from which follows

$$\|\tilde{D}^{-1} D\|_2 = \|(I + D^{-1} \Delta_D)^{-1} 2\| \leq \frac{1 - 2n\epsilon}{1 - 4n\epsilon} \quad (5.22)$$

and

$$\|\tilde{D}^{-1} \Delta_D\|_2 = \|(I + D^{-1} \Delta_D)^{-1} D^{-1} \Delta_D\|_2 \leq \frac{n\epsilon}{1 - 4n\epsilon}. \quad (5.23)$$

From Theorem 2.21, $\|\Delta_U\| \leq 3n^{3/2}\epsilon$, and it follows that

$$\|U^{-1} \Delta_U\|_2 \leq \|U^{-1}\|_2 \|\Delta_U\|_2 \leq 3n^3\epsilon \quad (5.24)$$

and thus

$$\|\tilde{U}^{-1} \Delta_U\|_2 \leq \|\tilde{U}^{-1}\|_2 \|\Delta_U\|_2 \leq \|(I + U^{-1} \Delta_U)^{-1}\|_2 \|U^{-1}\|_2 \|U\|_2 \leq \frac{3n^3\epsilon}{1 - 3n^3\epsilon} \quad (5.25)$$

From Theorem 2.21 we have

$$\|\Delta_L\|_2 \leq n^{1/2}\|\Delta_L\|_1 \leq \frac{n^{3/2}(4n-1)\epsilon_0}{3(1-2n\epsilon_0)}$$

where $\epsilon_0 = \frac{6\epsilon}{1-\epsilon}$. With (2.16) this gives

$$\|\Delta_L L^{-1}\|_2 \leq \|\Delta_L\|_2 \|L^{-1}\|_2 \leq \frac{n^3(4n-1)\epsilon_0}{3(1-2n\epsilon_0)} \quad (5.25)$$

and

$$\|\tilde{L}\|_2 \|L^{-1}\|_2 \leq \|L\|_2 \|L^{-1}\|_2 + \|\Delta_L\|_2 \|L^{-1}\|_2 \leq n^3 + \frac{n^3(4n-1)\epsilon_0}{3(1-2n\epsilon_0)} \quad (5.26)$$

Taking the absolute value of equation (5.20) and substituting the bounds (5.21), (5.22), (5.23), (5.25), and (5.26) gives

$$\begin{aligned} |\tilde{\lambda} - \lambda| \|\tilde{y}^* x\| &\leq |\lambda| \|\tilde{y}^*\|_2 \|x\|_2 \left(\frac{n^3(4n-1)\epsilon_0}{3(1-2n\epsilon_0)} \right) \\ &\quad + |\lambda| \|\tilde{y}^*\|_2 \|x\|_2 \left(n^3 + \frac{n^3(4n-1)\epsilon_0}{3(1-2n\epsilon_0)} \right) \left(\frac{2n\epsilon}{1-2n\epsilon} \right) \\ &\quad + |\tilde{\lambda}| \|\tilde{y}^*\|_2 \|x\|_2 \left(\frac{3n^3\epsilon}{1-3n^3\epsilon} \right) \left(\frac{1-2n\epsilon}{1-4n\epsilon} \right) \\ &\quad + |\tilde{\lambda}| \|\tilde{y}^*\|_2 \|x\|_2 \left(\frac{3n^3\epsilon}{1-3n^3\epsilon} \right) \left(\frac{n\epsilon}{1-4n\epsilon} \right) \end{aligned}$$

and thus

$$\begin{aligned} |\tilde{\lambda} - \lambda| &\leq |\lambda| \left(\frac{10n^4\epsilon_0 - n^3\epsilon_0 - 12n^5\epsilon^2}{3(1-2n\epsilon_0)^2} \right) \sec \theta(\tilde{y}, x) \\ &\quad + |\tilde{\lambda}| \left(\frac{3n^3\epsilon}{1-3n^3\epsilon} \right) \left(\frac{1-n\epsilon}{1-4n\epsilon} \right) \sec \theta(\tilde{y}, x) \end{aligned}$$

Use $|\tilde{\lambda}| \leq |\tilde{\lambda} - \lambda| + |\lambda|$ and rearrange to produce the desired result.

□

Thus, small perturbations in the diagonal dominant parts and the off diagonal entries lead to perturbations in the eigenvalues that are small multiples of a condition number of the eigenvalue. In Theorem 5.11, $\tilde{\lambda}$ can be any eigenvalue of \tilde{A} , but if $\tilde{\lambda}$ is not a good approximation of λ , then $\tilde{y}^* x$ will be small because $y^* x = 0$ if y a left eigenvector of A corresponding to some eigenvalue different from λ .

Copyright© Megan Dailey, 2013.

Chapter 6 Conclusions

We have systematically developed relative perturbation bounds for various linear algebra problems for diagonally dominant matrices. We have proved that small relative perturbations in the diagonally dominant parts and off diagonal entries of diagonally dominant matrices with nonnegative diagonals produce small relative perturbations in the LDU factorization. By using column diagonal dominance pivoting we ensure the L factor is column diagonally dominant, and thus well-conditioned. This allowed us to then prove similar perturbation bounds for the eigenvalues of nonsymmetric diagonally dominant matrices with nonnegative diagonals and for the singular values for general diagonally dominant matrices. The results relied upon perturbation results for the determinant which also led to relative perturbation for the inverse and solutions of linear systems involving diagonally dominant matrices and for the eigenvalues of symmetric indefinite diagonally dominant matrices.

In [43], an algorithm is presented that accurately computes all singular values with small relative error for diagonally dominant matrices. It is shown in [13, 43] that this algorithm also computes the L , D , and U factors of the LDU factorization of diagonally dominant matrices with small relative errors. Our results show that it is also possible to more accurately compute the inverse of a diagonally dominant matrix and eigenvalues of a nonsymmetric diagonally dominant matrix. For future works, we shall study algorithms for these problems.

Bibliography

- [1] Attahiru Sule Alfa, Jungong Xue, and Qiang Ye. Accurate computation of the smallest eigenvalue of a diagonally dominant m-matrix. *Mathematics of Computation*, 71(237):pp. 217–236, 2002.
- [2] Attahiru Sule Alfa, Jungong Xue, and Qiang Ye. Entrywise perturbation theory for diagonally dominant M-matrices with applications. *Numerische Mathematik*, 90(3):401–414, 2002.
- [3] Jesse Barlow and James Demmel. Computing accurate eigensystems of scaled diagonally dominant matrices. *SIAM Journal on Numerical Analysis*, 27(3):pp. 762–791, 1990.
- [4] Anders Barrlund. Perturbation bounds for the LDLH and LU decompositions. *BIT*, 31(2):358–363, June 1991.
- [5] F. Bauer and C. Fike. Norms and exclusion theorems. *Numerische Mathematik*, 2:137–141, 1960. 10.1007/BF01386217.
- [6] M.Isabel Bueno and Froilán M. Dopico. Stability and sensitivity of tridiagonal LU factorization without pivoting. *BIT Numerical Mathematics*, 44(4):651–673, 2004.
- [7] Megan Dailey, Froilán Dopico, and Qiang Ye. New relative perturbation bounds for LDU factorizations of diagonally dominant matrices. in progress.
- [8] Megan Dailey, Froilán Dopico, and Qiang Ye. Relative perturbation theory for diagonally dominant matrices. in progress.
- [9] James Demmel. *Applied Numerical Linear Algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
- [10] James Demmel, Ming Gu, Stanley Eisenstat, Ivan Slapnicar, Kresimir Veselic, and Zlatko Drmac. Computing the singular value decomposition with high relative accuracy. *Linear Algebra and its Applications*, 299(1-3):21 – 80, 1999.
- [11] James Demmel and W. Kahan. Accurate singular values of bidiagonal matrices. *SIAM Journal on Scientific and Statistical Computing*, 11:873–912, 1990.
- [12] James W. Demmel and William Gragg. On computing accurate singular values and eigenvalues of matrices with acyclic graphs. *Linear Algebra and its Applications*, 185(0):203 – 217, 1993.
- [13] Froilán Dopico and Plamen Koev. Perturbation theory for the LDU factorization and accurate computations for diagonally dominant matrices. *Numerische Mathematik*, 119:337–371, 2011. 10.1007/s00211-011-0382-3.

- [14] Froilán Dopico, Plamen Koev, and Juan Molera. Implicit standard Jacobi gives high relative accuracy. *Numerische Mathematik*, 113:519–553, 2009. 10.1007/s00211-009-0240-8.
- [15] Froilán M. Dopico and Plamen Koev. Accurate symmetric rank revealing and eigendecompositions of symmetric structured matrices. *SIAM Journal on Matrix Analysis & Applications*, 28(4):1126 – 1156, 2006.
- [16] Froilán M. Dopico and Juan M. Molera. Accurate solution of structured linear systems via rank-revealing decompositions. *IMA Journal of Numerical Analysis*, 32(3):1096–1116, 2012.
- [17] S. Eisenstat and I. Ipsen. Relative perturbation techniques for singular value problems. *SIAM J. Numerical Analysis*, 32(6):1972–1988, 1995.
- [18] S. Eisenstat and I. Ipsen. Relative perturbation results for eigenvalues and eigenvectors of diagonalisable matrices. *BIT Numerical Mathematics*, 38:502–509, 1998. 10.1007/BF02510256.
- [19] S. Eisenstat and I. Ipsen. Three absolute perturbation bounds for matrix eigenvalues imply relative bounds. *SIAM J. Matrix Analysis & Applications*, 20(1):149 – 158, 1998.
- [20] F.R. Gantmacher. *The Theory of Matrices*. Chelsea, New York, 1960.
- [21] Ji guang Sun. Componentwise perturbation bounds for some matrix decompositions. *BIT Numerical Mathematics*, 32(4):702–714, 1992.
- [22] Nicholas J. Higham. A survey of componentwise perturbation theory in numerical linear algebra. In *in Mathematics of Computation 1943–1993: A Half Century of Computational Mathematics*, pages 49–77, 1994.
- [23] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, 1996.
- [24] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [25] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [26] I. Ipsen. Relative perturbation results for matrix eigenvalues and singular values. *Acta Numerica*, 7(1):151–201, 1998.
- [27] P. Koev. Accurate eigenvalues and SVDs of totally nonnegative matrices. *SIAM Journal on Matrix Analysis and Applications*, 27(1):1–23, 2005.
- [28] Chi-Kwong Li and Ren-Cang Li. A note on eigenvalues of perturbed hermitian matrices. *Linear Algebra and its Applications*, 395(0):183 – 190, 2005.

- [29] R. Li. Relative perturbation theory. III. more bounds on eigenvalue variation. *Linear Algebra and its Applications*, 266:337 – 345, 1997.
- [30] Ren-Cang Li. A converse to the bauer-fike type theorem. *Linear Algebra and its Applications*, 109:167–178, 1988.
- [31] Ren-Cang Li. On perturbation theorems for the generalized eigenvalues of regular matrix pencils. *Chinese Journal of Numerical Mathematics and Applications*, 11:24–35, 1989.
- [32] Ren-Cang Li. Perturbation bounds for generalized eigenvalues I. *Chinese Journal of Numerical Mathematics and Applications*, 11:1–9, 1989.
- [33] Ren-Cang Li. Perturbation bounds for generalized eigenvalues II. *Chinese Journal of Numerical Mathematics and Applications*, 11:34–43, 1989.
- [34] Ren-Cang Li. Bounds on perturbations of generalized singular values and of associated subspaces. *SIAM Journal on Matrix Analysis and Applications*, 14:195–234, 1993.
- [35] Ren-Cang Li. Relative perturbation theory I: eigenvalue and singular value variations. *SIAM Journal on Matrix Analysis and Applications*, 19:956–982, 1998.
- [36] Ren-Cang Li. Relative perturbation theory II: eigenspace and singular space variations. *SIAM Journal on Matrix Analysis and Applications*, 20:471–492, 1999.
- [37] Ren-Cang Li, Yuji Nakatsukasa, Ninoslav Truhar, and Wei guo Wang. Perturbation of multiple eigenvalues of Hermitian matrices. *Linear Algebra and its Applications*, 437(1):202 – 213, 2012.
- [38] Ren-Cang Li and G.W. Stewart. A new relative perturbation theorem for singular subspaces. *Linear Algebra and its Applications*, 313(13):41 – 51, 2000.
- [39] J.M. Pena. LDU decompositions with L and U well conditioned. *Electronic Transactions on Numerical Analysis*, 18:198–208, 2004.
- [40] Daniel A. Spielman and Shang-Hua Teng. Solving sparse, symmetric, diagonally-dominant linear systems in time $O(m^{1.31})$. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '03, pages 416–, Washington, DC, USA, 2003. IEEE Computer Society.
- [41] Gilbert W. Stewart and Ji guang Sun. *Matrix Perturbation Theory*. Academic Press, 1990.
- [42] J. H. Wilkinson, editor. *The Algebraic Eigenvalue Problem*. Oxford University Press, Inc., New York, NY, USA, 1988.
- [43] Qiang Ye. Computing singular values of diagonally dominant matrices to high relative accuracy. *Mathematics of Computation*, 77(264):2195–2230, 2008.

- [44] Qiang Ye. Relative perturbation bounds for eigenvalues of symmetric positive definite diagonally dominant matrices. *SIAM Journal on Matrix Analysis & Applications*, 31(1):11 – 17, 2009.

Vita

Megan Dailey
Born in Frankfort, Kentucky

Education

- **University of Kentucky**, Lexington, Kentucky
 - M.A., Mathematics, May 2008
- **Centre College**, Danville, Kentucky
 - B.S., Mathematics and Chemistry, May 2006

Teaching Experience

- **Teaching assistant**
 - University of Kentucky, August 2006 - May 2013
- **Visiting Instructor**
 - Centre College, August 2011 - December 2012